

博士論文（要約）

**The Encoding and Maintenance of Time Intervals
in Visual Working Memory
and the Underlying Neural Mechanisms**

（視覚の作業記憶における間隔時間の符号化と維持
およびその神経機序の検証）

範 志偉 (FAN ZHIWEI)

1	General introduction	1
1.1	From time dimension to temporal items	1
1.2	The cognitive and neural mechanisms of time perception	6
1.3	Entrainment of neural oscillations and its effect on time perception	10
1.4	Neural entrainment induced by tACS	10
1.5	Thesis scope	11
2	Study 1: Time intervals are represented in working memory as discrete items	12
2.1	Study 1: Introduction	12
2.2	Study 1: Experiment 1	14
2.3	Study 1: Experiment 2	27
2.4	Study 1: Experiment 3	35
2.5	Study 1: General discussion	42
3	Study 2: Periodic events modulate encoding of time intervals in working memory	46
3.1	Study 2: Introduction	46
3.2	Study 2: General methods	49
3.3	Study 2: Experiment 1	52
3.4	Study 2: Experiment 2	55
3.5	Study 2: Experiment 3	57
3.6	Study 2: Experiment 4	60
3.7	Study 2: General discussion	63
4	Study 3: Oscillatory mechanisms underlying working memory of time intervals	66
4.1	Study 3: Introduction	66

4.2	Study 3: Methods	68
4.3	Study 3: Results	73
4.4	Study 3: Discussions	75
5	General discussion	78
5.1	Time dimension <i>of</i> the mind can be represented as temporal items <i>in</i> the mind	78
5.2	Time encoding depends on external events and internal oscillations	81
5.3	Neural oscillatory mechanisms underlying time encoding	83
5.4	Short-term tACS can induce entrainment in specific stages of processing	85
5.5	Conclusion	86
6	Reference	88
7	Acknowledgement	101

1 General introduction

Time is an intuitive concept, and we are familiar with the experience of time. We may imagine time as a unidirectional flow, think of time as a basic dimension of the universe, or use time as a reference to describe life events. However, daily experiences do not provide insight into the mechanisms underlying time in the mind. Is time also a dimension of the mind? Can it be represented in the mind in a manner similar to that of other sensory information such as vision and audition? What are the cognitive and neural mechanisms underlying the representation of time in the mind? These questions are the focus of many researchers that study psychological time.

As a topic in psychological science, the psychology of time has been relatively less attended to, compared to that of space. Over the past few decades, psychological time has received a renaissance of interest. Time has now become a major topic in the fields of psychology and neuroscience (Hancock & Block, 2012). Psychological time consists of three major aspects: succession (order or sequence), temporal perspective (past, present, and future), and time interval/duration which will be the focus of this thesis (Block, 1990).

1.1 From time dimension to temporal items

1.1.1 Time may not be an independent dimension

In physics, time is not considered an independent dimension as space and does not have equal status to space. But In psychology and neuroscience, time and space remain independent coordinates of experience. People in most cultures use time as a basic dimension to describe the universe as well as life events. In some cultures, such as the Amondawa culture of Amazonia, however, people do not conceive of time as something independent, in which events occur (Sinha, Sinha, Zinken, & Sampaio, 2011). Therefore, time is not a universal dimension of the mind which is independent of events (Buzsáki & Llinás, 2017).

Instead, perceived time is affected by events. For example, even with the same travel-time, the return trip would be felt shorter than the outward trip (Ozawa, Fujii, & Kouzaki, 2015). Routes of the same travel-time may be perceived differently. This may be due to the different events happening on the routes and thus yielding the different psychological states such as familiarity with the routes (Jafarpour & Spiers, 2017).

1.1.2 Time can be content in the human mind

Not all societies use calendars and clock systems to quantify time (Sinha et al., 2011). However, early researchers in the field of psychological time have proposed models with an internal clock system for measuring time and guiding the timing of behavior (Church, 1984; M. Treisman, 1963). Recently, there has been a consensus that time is not necessarily a dimension for encoding experience; instead, it is encoded through the encoding of experience (Buzsáki & Llinás, 2017; Tsao et al., 2018). Regardless of whether time is truly a dimension or not, a critical question is how time transfers from the so-called dimension of the human mind to the content in the human mind (van Wassenhove, 2017). In more psychological terms, this is referred to as the perception and memory of time.

If time is encoded through the encoding of experience which contains several events, then time is encoded through the processing of events in at least two steps (Johnston & Nishida, 2001): first, to process a number of events; second, to extract the temporal aspects (in addition to visual, auditory, and other aspects) or temporal contents of the events. Temporal contents of events are then stored as part of event-related memories (Block, 1990). Indeed, Gibson (1975) claimed, "there is no such thing as the perception of time, but only the perception of events and locomotion". Time dimension, which has traditionally been assumed to be continuous, may become several discrete temporal contents via the encoding of events; these contents may be encoded, maintained, and retrieved.

1.1.3 Interval timing perception and working memory

Interval timing perception is defined as the perception of time intervals in the range of hundreds of milliseconds to minutes (Buhusi & Meck, 2005; Kononowicz, van Rijn, & Meck, 2018). Models of interval timing perception such as time discrimination models explain time perception by parsing it into several stages of information-processing. These include clock, storage, and decision/comparison stages (Matell & Meck, 2004); in other words, encoding, maintenance, and retrieval/decision stages (Kulashekhar, Pekkola, Palva, & Palva, 2016). In information-processing (IP) models of interval timing perception, working memory has been considered an essential process (Gibbon, Church, & Meck, 1984; M. Treisman, 1963). Working memory, more than short-term memory (Baddeley & Hitch, 1974), is the online workspace for actively maintaining and manipulating relevant information (Baddeley, 2003). There are different types of working memory which correspond to specific modalities, such as

auditory (Delogu, Nijboer, & Postma, 2012) and visual working memory (Luck & Vogel, 1997). Visual working memory actively maintains visual information of events for several seconds to be used during ongoing tasks (Luck & Vogel, 2013). It can be functionally divided into several dissociable processing stages such as encoding, maintenance, and retrieval/decision (Pearson, Raskevicius, Bays, Husain, & Hospital, 2014; Woodman & Vogel, 2005).

With regard to neural mechanisms, an integrative model proposed that interval timing perception and working memory originate from the same neural oscillatory mechanisms for processing events which also define time intervals (Gu, van Rijn, & Meck, 2015). As described in 1.1.2, time intervals are encoded through the encoding of events where they are extracted as the temporal aspects of the events. There are other aspects which can also be extracted, such as the visual aspects (texture, color, shape, etc.) of events. The same neural oscillatory activity of the same neuronal populations for the same events is hypothesized to encode visual aspects (such as visual textures) in working memory, as well as temporal aspects (such as time intervals) in interval timing perception. Thus, these two processes differ only in terms of the dimension of neural oscillations that is utilized for the extraction of visual or temporal aspects of the same events (Gu et al., 2015). In this thesis, perception and working memory of time intervals are considered two sides of the same coin. Working memory of time intervals shares the same processing stages with interval timing perception, including encoding, maintenance, and retrieval/decision stages of time interval processing (Gu et al., 2015; Kulashekhar et al., 2016), as shown in Figure 1.1.

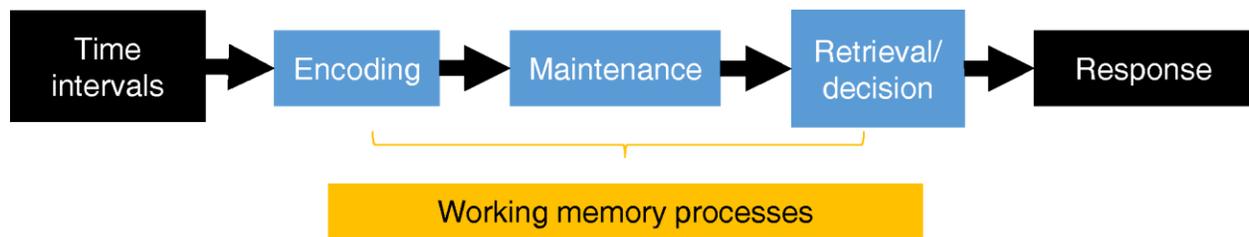


Figure 1.1. Working memory processes for time intervals. Time intervals are subjected to the encoding, maintenance, and retrieval/decision stages of processing, and finally utilized for responses.

1.1.4 Representation of time intervals in visual working memory

It is still unclear how time intervals are stored in and retrieved from working memory. Previous studies on time have mainly focused on the encoding of single time intervals (Grondin, 2010; Wearden, 2016). Only a few studies have examined working memory of multiple time intervals. These studies have investigated working memory for multiple time intervals of auditory events (Manohar & Husain, 2016; Teki & Griffiths, 2014, 2016). However, these studies used the time reproduction task, yielding responses as a continuous variable, whereby the time interval is reproduced as a quantity rather than recognizing it as an object or item. Time was quantitatively measured as a continuous parameter rather than being qualitatively recognized as discrete items in these studies, thereby limiting the claims of these studies that time intervals can be stored as distinct items in working memory. Therefore, whether and how a continuous time dimension becomes discrete temporal items remains unclear. Furthermore, working memory for time intervals has not been directly compared with working memory for non-temporal (for example, visual) modalities. Thus, the nature of working memory for time intervals remains unknown.

Considering visual working memory for multiple objects/items, the Noisy Exemplar Model (NEMo) of memory (Kahana & Sekuler, 2002) proposed that visual objects are represented as noisy coordinates in a multi-dimensional feature space (see Figure 1.2). Visual objects consist of features (such as color, orientation, and spatial frequency) which define the objects or serve as the dimensions to represent the objects (Blaser, Pylyshyn, & Holcombe, 2000; A. Treisman & Paterson, 1984). In addition, from either the perspective of NEMo or the signal detection theory, because of the inevitable variability from external inputs or internal representations, each object is represented as a multivariate normal distribution with noise (Kahana & Sekuler, 2002; Macmillan & Creelman, 2005). Thus, visual objects are represented as separate but noisy items in working memory. As shown in Figure 1.2, during visual recognition, visual objects are presented one after another as study items, and another object is presented as the probe. The recognition judgment of whether the probe is one of the study items depends on the distance or the similarity between the probe and study items. Furthermore, the noise of each item increases as the time goes by. Thus S1 should be noisier than S2 and S3, and S2 should be noisier than S3, and so on. The recognition judgment is affected by their presentation orders or the so-called serial positions.

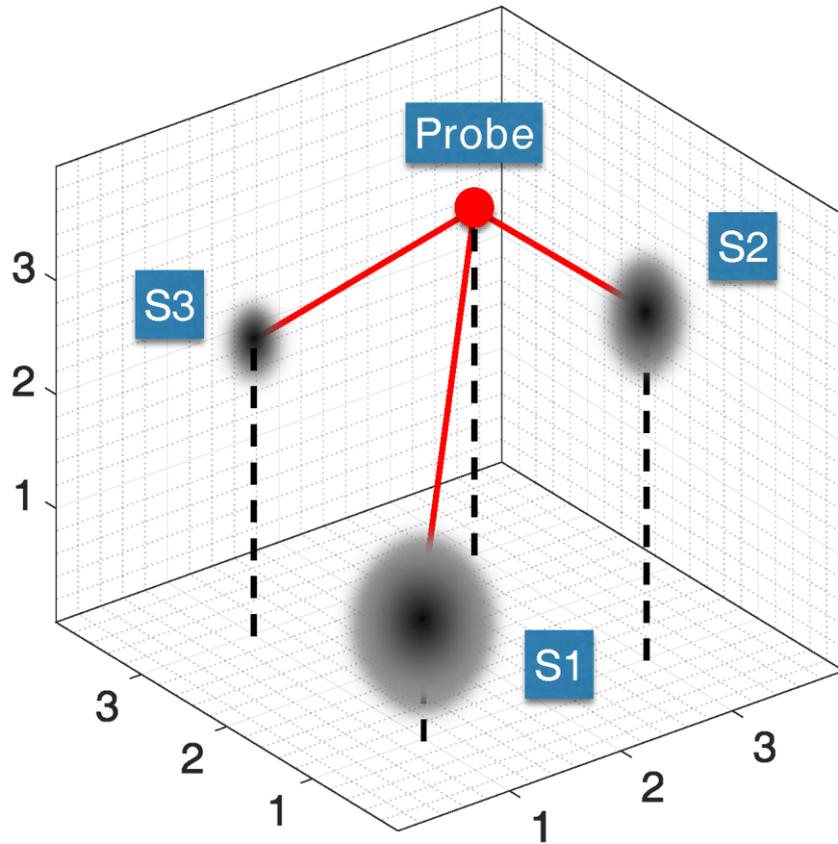


Figure 1.2. The three-dimensional feature space for the representation of items according to the Nosy Exemplar Model (NEMo), adapted from (Kahana & Sekuler, 2002). The dimensions represent the features such as spatial frequency (in 1, 2, and 3 cycles per degree), which define the items. Here assume three spatial frequencies, and therefore there are three dimensions. The three fuzzy ellipses are the three study items (S1, S2, and S3) with noise, which have been studied and remembered, and are to be recognized. The blurred area represents the noise associated with the encoding of the stimulus. The size of the blur corresponds to the size of the noise. The red point is the probe item. Recognition judgment involves judging whether the probe item is one of the study items. The judgments are based on the similarity (or distance), indicated by the red lines, between the probe and study items.

Investigation of whether temporal items are represented in the same manner as that of visual items may reveal not just the nature of working memory for time intervals, but may also provide extensive support for the NEMo representation of items.

1.2 The cognitive and neural mechanisms of time perception

Time is an abstract entity which is not directly observable. A “model is a description that helps us understand a phenomenon or process that is not directly observable (Block, 1990)”. For this reason, many models have been proposed specifically for psychological time and its cognitive and neural mechanisms. There are two main streams of time models: dedicated models which assume a special system providing a clock-like time reference and intrinsic models which do not make this assumption.

1.2.1 Intrinsic models and dedicated models

Intrinsic models, mainly put forth by Buonomano’s group, propose that timing may be derived from recognizing neural states without the need of an internal clock (Karmarkar & Buonomano, 2007; Paton & Buonomano, 2018). If a neural network is able to recognize activated states and reproduce the trajectories of the states across time and space, then it is capable of marking time intervals and thereby guiding the timing of behavior. Therefore, timing arises from the intrinsic properties of neural circuits per se. These models have gained reasonable traction, especially for time interval perception in the sub-second range (Buonomano, 2014).

Conversely, dedicated models assume a dedicated timing system for time intervals ranging from hundreds of milliseconds to minutes, imitating clocks used in daily life. These interval clock models include earlier models for humans (M. Treisman, 1963, 1984) and animals (Church, 1984), which were based on an information-processing perspective (Kurotsu, Marahiel, Muller, & Kleinkauf, 1982). However, internal clock models are limited and may be oversimplified and mechanistic, although later efforts were made to improve them. For example, attention was introduced as a component (Zakay & Block, 1997). Wittmann (2016) even stated, “The pacemaker-counter model is just a metaphor, anyway – a ticking stopwatch that can be turned on and off”.

The most modern and neurobiologically feasible dedicated model is the Striatal Beat-Frequency (SBF) model which proposes that cortico-striatal circuits may be the neural basis of a dedicated timing system, with

cortical neurons serving as oscillators and medium spiny neurons (MSNs) in the striatum (the main component of the basal ganglia, BG) acting as integrators and coincidence detectors (Buhusi & Meck, 2005; Matell & Meck, 2004). The encoding of the time intervals of an event is based on the coincident detection of neural oscillations at the conclusion of that event (see Figure 1.3).

In summary, these modeling efforts have demonstrated the complexity of time in the mind and highlighted that the mental representation of time is underscored by different levels of neuronal systems, from single neurons to subcortical-cortical networks. For example, a recent study reported the discovery of “time neurons” in the lateral entorhinal cortex, but rejected the idea of a constant internal clock because the activity of these neurons is affected by the experience of events (Tsao et al., 2018). The interaction between cortical neurons and time neurons in the lateral entorhinal cortex is not yet clear, but suggests a hierarchical network underlying the timing of behavior which may unite dedicated and intrinsic systems at different neuronal levels. For either stream of models, events are the basis for time encoding, either by defining time intervals, or by changing neural states (for intrinsic models) or neural oscillations (for the SBF model), thereby determining time in the mind.

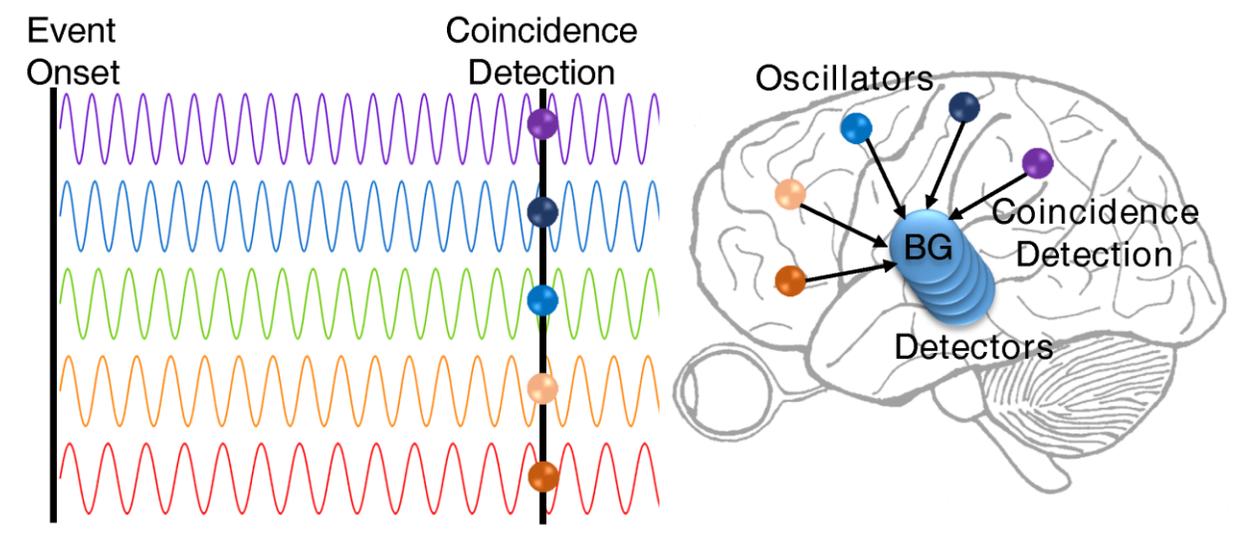


Figure 1.3. The SBF model with coincidence detection for the encoding of time intervals, adapted from Buhusi & Meck (2005). Left panel shows the oscillatory activity of cortical neurons which serve as oscillators. Right panel shows cortical neurons projecting onto MSNs in the Basal Ganglia (BG), which act as the detectors. Interval timing perception is based on the coincidental activation of the MSN detectors by the cortical neural oscillators (the five

colorful circles). For the to-be-timed interval of event, the phases of the oscillators are resting at the onset of the event (the first black line). At the end of the event (the second black line), the coincident activation pattern of the oscillators is detected by the detector tuned to that specific interval. Thus, the time interval is encoded as the readout of the detector.

1.2.2 **The role of neural oscillations**

Neural oscillations are ubiquitous in the brain at various frequency bands including delta, theta, alpha, beta, and gamma bands (Buzsáki & Draguhn, 2004) and are closely related to cognition (Wang, 2010), including interval timing perception (Gu, Kukreja, & Meck, 2018; Gu et al., 2015; Kulashekhar et al., 2016). Neural oscillations could be a possible neurobiological basis for oscillators in the SBF model. An oscillator may be a single neuron and an ensemble of neurons firing at the same oscillation frequency. Studies have provided evidence for different frequency bands of neural oscillations underlying different working memory processes for time intervals, such as gamma bands for encoding, beta bands for estimation, and theta to alpha bands for working memory maintenance and retrieval (Kononowicz & van Rijn, 2015; Kulashekhar, 2017; Kulashekhar et al., 2016; Wiener, Parikh, Krakow, & Coslett, 2018). However, research on the role of neural oscillations in time processing is still preliminary.

Instead of using the oscillation periods, the summation of periods, or the number of periods of oscillators to encode time intervals, the SBF model utilizes the coincidence detection of a group of oscillators with different frequencies. However, there is no simple relationship between oscillation frequency and encoded time (Matell & Meck, 2004; M. Treisman, 1963, 1984). It may not be the frequency but rather the synchronization (earlier or later) of neural oscillations that affects the encoding of time intervals. Therefore, any events that affect the synchronization of oscillations should affect time encoding.

1.2.3 **Neural oscillations modulated by periodic events affect time encoding**

Visual stimuli with temporally modulated luminance, known as visual flickers, have been used as events that affect neural oscillations. Visual flickers induce robust overestimation of time (Hashimoto & Yotsumoto, 2015, 2018; Herbst, Javadi, van der Meer, & Busch, 2013; Kanai, Paffen, Hogendoorn, & Verstraten, 2006; Ortega &

López, 2008; M. Treisman & Brogan, 1992). Flicker-induced time overestimation may be due to the periodic property of visual flickers which may modulate neural oscillations in the brain (Hashimoto & Yotsumoto, 2015, 2018). However, visual flickers are a distinct class of visual input and are unsuitable to be used for studying the general nature of visual working memory. In order to study visual working memory for stimuli other than visual flickers while modulating neural oscillations, other periodic stimulations should be used.

Non-invasive brain stimulation (NIBS) techniques, such as transcranial alternating current stimulation (tACS), can directly modulate ongoing neural oscillations to study the causal relationships between neural oscillations and cognitive functions (Polanía, Nitsche, & Ruff, 2018; Vosskuhl, Strüber, & Herrmann, 2018). Both periodic visual stimuli and electric stimulation may influence the encoding of time intervals by interfering with brain oscillations (see Figure 1.4). Note that sensory (e.g., visual) events also cause sensory and perceptual effects, such as the flickering perception caused by visual flickers. It is not known whether the direct modulation of brain oscillations depends on the sensory property of events. For instance, it is unclear whether the modulatory effect of tACS differs between flickering and static visual stimuli.

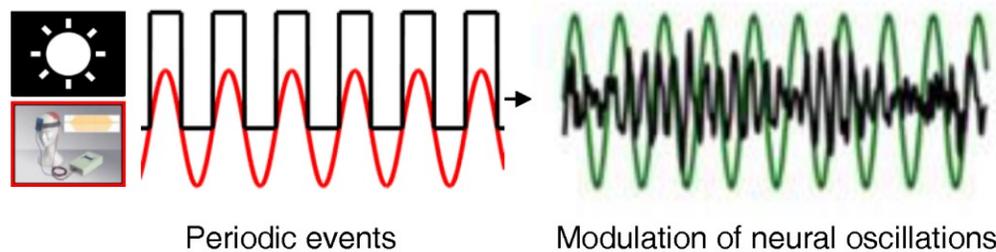


Figure 1.4. Periodic events such as visual flickers and tACS can modulate neural oscillations in the brain. In the upper left panel, the white disk with small white rectangles represents visual flickers. The stimulation pattern of visual flickers can be represented as a square wave (black wave in the middle panel). The picture in the lower left panel indicates the equipment that produces tACS. tACS applies a sinusoidal signal (red wave in the middle panel). The right panel shows neuronal oscillations (in black) modulated by a driving frequency (in green).

1.3 Entrainment of neural oscillations and its effect on time perception

Entrainment of brain oscillations refers to the temporary synchronization or resonance of brain oscillations to external periodic events (alternating current, magnetic pulses, visual flickers, and auditory tones) (Pikovsky, Rosenblum, & Kurths, 2003; Thut, Schyns, & Gross, 2011; Vossen, Gross, & Thut, 2015). The SBF model suggests that entrainment of oscillators should change time perception which is supported by studies of visual flickers-induced time dilation (Hashimoto & Yotsumoto, 2015, 2018). However, there are few studies on the effects of NIBS entrainment on the perception of time intervals. A recent study used tACS at beta frequency and found a role of beta oscillations in modulating the reference interval, thereby leading to longer judgment of the test interval (Wiener et al., 2018). However, the effect of tACS on the overestimation of time may be due to neural plasticity resulting from the long-term (20 minutes) continuous tACS used in the study or the combination of neural plasticity and neural entrainment, rather than neural entrainment alone. Moreover, the effects of neural entrainment on other properties of time perception, such as scalar property, have yet to be studied.

1.4 Neural entrainment induced by tACS

It has been 10 years since the first tACS study was published (Antal et al., 2008), yet the field of oscillatory brain stimulation is still in its infancy. Although oscillatory stimulation and neural entrainment by tACS have been improved in many ways, there are no reports on neural entrainment induced by short-term tACS (within or less than the timescale of seconds) and its effects on cognitive functions including the perception and memory of time (Strüber, Rach, Neuling, & Herrmann, 2015; Vosskuhl et al., 2018). There is evidence that short-term tACS of 1 s is too short to induce synaptic plasticity. Thus, short-term tACS may enable the separation of neural entrainment from synaptic plasticity, both of which are suggested to mediate the behavioral effects of tACS (Vosskuhl et al., 2018). In this thesis, I investigated the visual working memory of time intervals. By definition, visual working memory representation is maintained in sustained neural activity rather than in synaptic strength (Luck & Vogel, 2013). Visual working memory consists of several stages of processes. In this thesis, I mainly focused on the encoding stage. In this regard, short-term tACS may be a useful tool to facilitate the investigation by focusing on a specific stage and avoiding the induction of synaptic plasticity.

1.5 Thesis scope

This thesis goes beyond previous studies which have mainly focused on a single time interval and used corresponding traditional investigation methods. As multiple time intervals are essential for processing complex events, it is imperative to understand how multiple time intervals are encoded, maintained, retrieved, and utilized in decision-making for future actions. Surprisingly, there are limited studies on the representation of multiple time intervals (Manohar & Husain, 2016; Teki & Griffiths, 2014, 2016).

In Study 1, I investigated working memory representation of multiple time intervals and compared it with that for multiple visual textures in a recognition task. The results of Study 1 showed the typical working memory characteristics for time intervals, similar to those for visual textures, and demonstrated that the so-called time dimension is represented as discrete but noisy temporal items in working memory.

In Study 2, I examined how time intervals are encoded as individual items. As mentioned in section 1.1.2, time intervals should be encoded from single events, and affected by event properties, especially the periodic dynamics, considering the possible neural oscillatory mechanisms underlying the encoding of time intervals. I used visual flickers and tACS to investigate the effects of periodic events on time encoding in a reproduction task. The results of Study 2 indicated periodic simulations (visual flickers) could increase the reproduced duration; both visual flickers and tACS reduced the coefficient of variation (CV) which corresponded to the noise of the temporal items in the NEMO-based working memory space. I observed that the effects of periodic events such as visual flickers and tACS may modulate neural oscillations and result in altering time encoding.

In Study 3, I assessed the effects of tACS on other aspects of working memory performance for temporal items and the underlying neural mechanisms. However, the results of Study 3 failed to show significant effects of tACS on behavioral performance and did not provide EEG evidence for short-term tACS-induced neural entrainment either. These results suggest that the effects of short-term tACS may depend on the nature of the tasks and different cognitive processes involved, or the entrainment might have been too weak.

2 Study 1: Time intervals are represented in working memory as discrete items

2.1 Study 1: Introduction

As an event persists, the time interval defined by the event needs to be constantly updated in working memory. In some cases, especially for human activities, it may be necessary to remember several time intervals of events simultaneously in working memory. However, previous studies on time have mainly focused on the perception of single time intervals, and for this reason, they adopted measures such as verbal estimation, reproduction, production, and comparison (Grondin, 2010; Wearden, 2016).

In the present study, we used Sternberg's item recognition task (Sternberg, 1966) to examine working memory for more than one time intervals, and compared its characteristics with those for visual objects. Sternberg's item recognition task is a widely used paradigm for studying working memory for items. In the task, several items (study items) are sequentially presented and remembered, and then they are compared with a final item (probe) after a short delay. The participants are expected to indicate whether the probe is a target (the same as one of the study items) or a lure (not the same as any of the study items). The working memory performance on the recognition task can be explained using the signal detection theory (SDT) (Macmillan & Creelman, 2005). The SDT explains how well the items are perceptually discriminated between targets and lures, measured by sensitivity (d'), and what criterion is used to determine if a particular item is the target or the lure, measured by decision criterion (C). The d' reflects a relatively earlier (for example, sensory) stage of processing, and the C reflects a later (decision) stage of processing. Modeling item recognition using SDT allows independent assessments of the two stages of processing (Rotello, 2017). When the SDT is incorporated into the recognition task, a receiver operating characteristic (ROC) curve can be derived from plotting hit rate (the proportion of correct recognition of targets) as a function of false-alarm rate (the proportion of recognizing lures as a target). The ROC curve reflects a constant d' that represents discriminability and a variable C that represents response bias.

One key characteristic variable of working memory in Sternberg's task is memory load (or set size, list length) (Sternberg, 1966). Specifically, the memory performance worsens with an increase in the memory load, that is, an increase in the number of study items. This memory load effect indicates a limited working memory capacity for items. For visual items, there are two competing views on working memory capacity. One perspective proposes

that items may be stored in a fixed number of slots with a fixed resolution (Luck & Vogel, 1997; Zhang & Luck, 2008), whereas the other proposes that the working memory resource can be flexibly distributed across all items (Bays & Husain, 2008; Ma, Husain, & Bays, 2014). Despite the difference, both views are related to the discrete item-based representations (Xie & Zhang, 2017), that is, both views at least agree that visual objects are represented as discrete items in visual working memory, and that memory load influences the memory performance for visual items. Therefore, if time intervals are represented as items similar to the visual objects, there should also be a memory load effect for the temporal items. To clarify, the term “discrete” used in this study corresponds to the item representation, that is, being represented individually and recognizable, different from what is often used to describe the “slot” property of working memory storage (Luck & Vogel, 1997; Zhang & Luck, 2008).

Another prominent characteristic variable for visual working memory is the serial position of an item on the study list (Oberauer, 2003). The serial position effect refers to the better recall or recognition of the first or last presented items than for the items presented in the middle. This effect contains two types: one is the primacy effect, (the performance is better for the first or first few items presented), and the other is the recency effect (the performance is better for the last or last few items presented). Note that the recognition of visual items often yields no primacy effect, but only exhibits a last-item recency effect (for example, Avons, 1998; Phillips & Christie, 1977).

The similarity between items is also an important variable in visual working memory for multiple items (Jiang, Lee, Asaad, & Remington, 2016). According to NEMo (Kahana & Sekuler, 2002), similarity is defined as the reciprocal of the Euclidian distance between two items in this multi-dimensional space. For a probe that does not match any of the study items, the larger the sum of similarities between the probe and each of the study items, the higher is the possibility that it would wrongly be judged as the target. Whether the probe is judged as the target or the lure is determined by whether the summed similarity crosses a decision criterion (Kahana & Sekuler, 2002; McKinley & Nosofsky, 1996; Nosofsky, 1984). The summed similarity requires the discrete representations of items before being computed (F. Zhou, Kahana, & Sekuler, 2004). If time intervals are represented discretely, the judgment of lure intervals should be subject to the summed similarity as well.

Further, the item representations of visual objects are related to binding (Burwick, 2014). Binding integrates different visual features (for example, color and shape) into a distinctive object (for example, a red triangle) so that they can be recognized as a distinctive item. Therefore, if time intervals are represented as items, the

temporal features may be subject to binding. A recent study provided evidence suggesting that time intervals can be bound into auditory event representations (Bogon, Thomaschke, & Dreisbach, 2017). On the other hand, however, it remains unclear if the recognition of temporal features is also affected by visual features.

In the present study, we hypothesized that similarly to visual objects, time intervals can be stored as discrete items in working memory. The presence of the effects of memory load, serial position, and similarity for time intervals would provide support for this hypothesis. The possible binding of temporal and visual features, which may be the underlying mechanism of item representations of time intervals, was also investigated. In addition, previous studies suggested that time intervals are more modality specific for the sub- than for the supra- second range (Mioni et al., 2016; Rammsayer, Bortner, & Troche, 2015). The processing of these two ranges of time intervals may be influenced by visual features in different ways. Therefore, we also compared working memory for time intervals in sub- and supra-second ranges.

2.2 Study 1: Experiment 1

In Experiment 1, we measured and compared working memory for visual textures and for intervals in the sub-second range. The effects of memory load, serial position, and similarity were examined to evaluate the hypothesis.

2.2.1 Materials and Methods

2.2.1.1 Participants

Sixteen naïve volunteers from The University of Tokyo (7 females, mean age: 22 years, range: 19–28 years) participated in Experiment 1, after excluding one participant whose average accuracy of recognition did not exceed the chance level.

In all experiments, the participants gave written informed consent to participate in the experiment in accordance with the Declaration of Helsinki. The protocol was approved by the Institutional Review Boards of the University of Tokyo, and all experiments were carried out in accordance with the guidelines set by the Ethics Committee of the University of Tokyo. All participants reported having normal or corrected-to-normal vision.

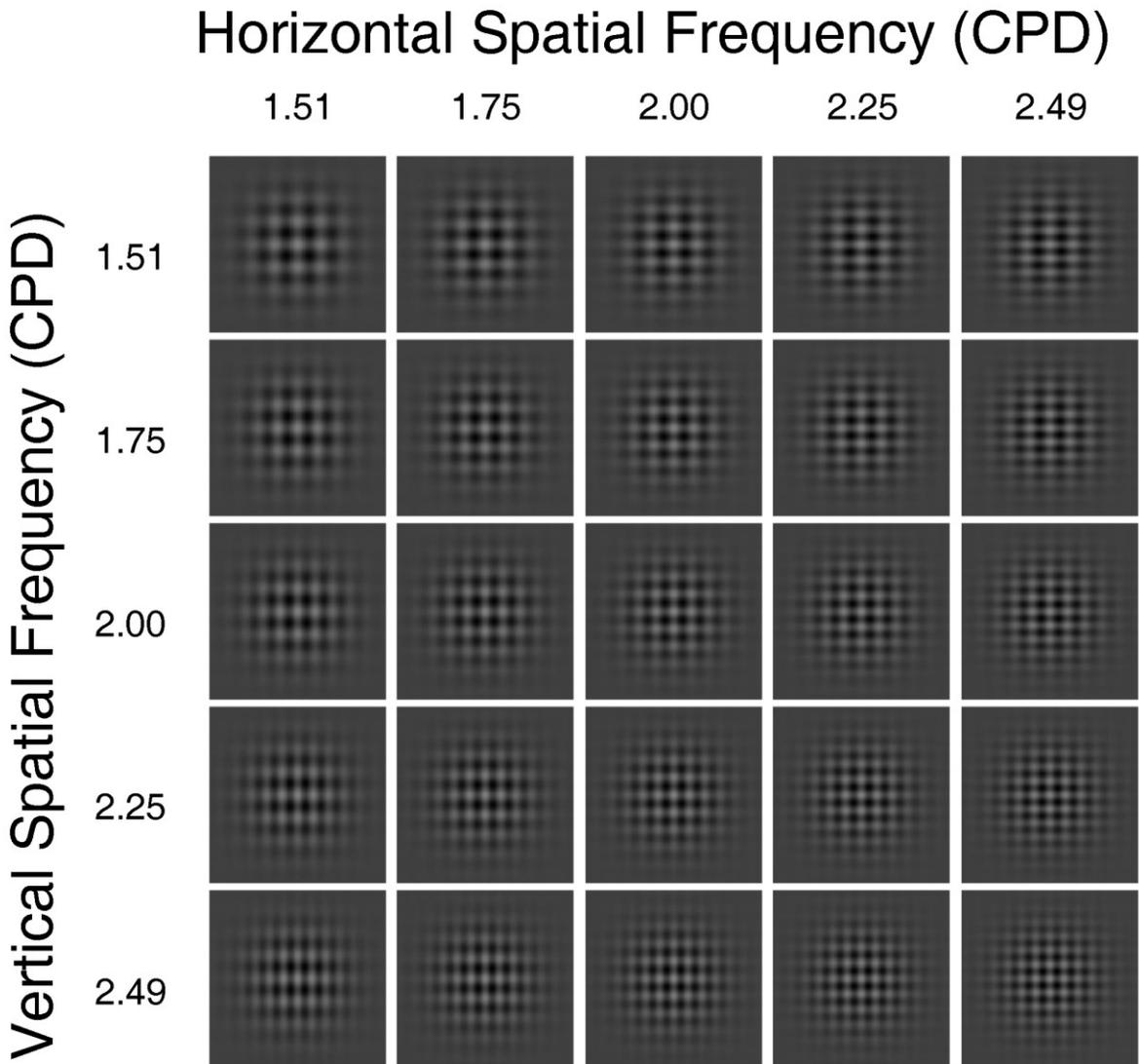


Figure 2.1 Gratings used in Experiment 1~2, which were of 25 different 2D-spatial frequencies derived by combining 5 vertical and 5 horizontal spatial frequencies of 1.51, 1.75, 2, 2.25, and 2.49 (CPD). The row shows the horizontal spatial frequency and the column shows the vertical spatial frequency.

Stimuli for each trial were drawn from a pool of two-dimensional (2D) sinusoidal gratings that are often employed in studying recognition/working memory (for example, Kahana & Sekuler, 2002; Yotsumforo, Kahana,

McLaughlin, & Sekuler, 2008). Gratings were of 25 different spatial frequencies, derived by combining five vertical and five horizontal frequencies of 1.51, 1.75, 2, 2.25, and 2.49 cycles per degree (CPD), similar to those used by Yotsumoto et al. (2008) (see Figure 2.1). The gratings were presented at 10 different time intervals from 0.25 to 0.97 s, with a linear increment of 0.08 s. The numbers of different visual or temporal stimuli rendered the probe prediction impossible. The luminance profile of the gratings was as follows:

$$L(x, y) = L_{avg} * \left(1 + c * \left(\cos(2 * \pi * f_x * x) + \cos(2 * \pi * f_y * y)\right) / 2\right),$$

where L_{avg} is the mean luminance, of 20 cd/m²; c is the contrast; f_x is the spatial frequency of the vertical fundamental component, in CPD; and f_y is the spatial frequency of the horizontal fundamental component. The contrast was set at 1, which was well above the detection threshold. The gratings were five degrees wide and they were windowed by a 2D Gaussian function with a space constant of one degree. The stimulus parameters were determined with an intention to equalize task difficulties between the interval session and the visual texture session.

Working memory performance for time intervals and for visual textures was measured in two separate sessions. In the interval session, participants judged the time intervals while ignoring the visual textures, and in the texture session, participants judged the visual textures while ignoring the time intervals. Identical sets of stimuli were used in both sessions. For each trial in each session, the study items were either one, two, or three gratings, yielding three memory loads. The spatial frequencies of the study items were randomly selected from the 25 different spatial frequencies forming different visual textures, and the time intervals of the study items were randomly selected from the 10 different time intervals. Another item was drawn as the probe, with the following constraints: in a target trial, the probe matched one of the study items, and the serial position in each memory load was probed equally frequently; and in a lure trial, the probe did not match any of the study items. The lure probe was also set such that it was not the neighbor of any of the study items in terms of magnitude, for either visual textures or time intervals. This was to control the similarity between the lure probes and the study items and to make the lure probes distinguishable from the study items. In the interval session where participants judged the time intervals while ignoring the visual textures, the visual textures served as context features. In the texture session where participants judged the textures while ignoring the time intervals, the time intervals served as context features. When

the context feature of the probe matched the context feature of the study item, the context was considered to be “repetition,” and when the context feature did not match the study item, the context was considered to be “switch.”

2.2.1.3 *Apparatus*

Stimuli were generated on a desktop computer (HP Compaq 8200 Elite), using MATLAB 2014a (The MathWorks Inc., Natick, MA, USA) with Psychophysics Toolbox Version 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), and they were presented on a 22-inch CRT monitor (Diamondtron M² RDF223H, MITSUBISHI, Tokyo, Japan) with a refresh rate of 85 Hz and a resolution of 1024 × 768 pixels. The monitor was gamma calibrated by a ColorCAL MKII colorimeter (Cambridge Research Systems Ltd, Rochester, Kent, UK). The experiment was conducted in a dark room. Participants sat in front of the monitor, at a distance of 57.3 cm, with their heads on a chin rest.

2.2.1.4 *Procedure*

All participants completed two sessions. The order of the sessions was counterbalanced across participants. Participants were instructed to refrain from verbally counting (Rattat & Droit-Volet, 2012). Figure 2.2 shows the time course of a typical trial in a session. Each trial began with a black fixation cross presented at the center of the screen for 0.3–0.5 s, followed by N (1, 2, or 3) gratings with different visual textures and time intervals, consecutively presented as study items. If N > 1, the black fixation cross appeared again for 0.3–0.5 s as an inter-stimulus interval (ISI), followed by another study item. During the presentation of the study items, the fixation cross remained black. The cross then turned red for 1.0–1.2 s, and finally, the probe grating appeared. After the probe was presented, the cross turned blue. Participants were asked to answer whether the probe matched one of the study items by pressing a button assigned to “Yes” or “No.” The right and left arrow keys were used to register the responses. The participants used their right index and middle fingers to make the responses. For half of the participants, the right and left arrow keys were assigned to “Yes” and “No,” respectively, while the keys were reversed for the other half of the participants. Additionally, participants were asked to respond as quickly as possible, without sacrificing accuracy. After the response was registered, the trial was ended and it was followed by a new trial after an inter-trial interval (ITI) of 0.7 s.

Each session consisted of 24 practice trials and 360 test trials composed of a combination of factors as memory load (number of study items: 1, 2, or 3), serial position (1 for a memory load of 1; 1 or 2 for a memory load of 2; 1, 2, or 3 for a memory load of 3), context (repetition or switch) and trial type (target trials or lure trials).

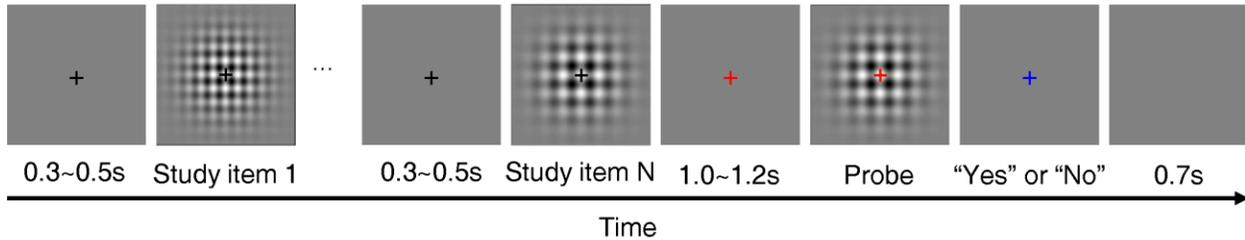


Figure 2.2. Time course of the recognition/working memory task in Experiments 1 and 3. The number of study items (N) was either one, two, or three in Experiment 1, and was always three in Experiment 3.

2.2.2 Results

Individual trials associated with response times (RTs) beyond three standard deviations from the mean RT were removed from further analyses. The same exclusion criteria were also employed in Experiments 2 and 3. The average accuracy across all available participants for the time intervals was $65.93\% \pm 5.70\%$, for the visual textures was $72.96\% \pm 6.69\%$; the group RT across all available participants for the time intervals was 0.86 ± 0.39 s, and for the visual textures was 0.58 ± 0.24 s.

2.2.2.1 Memory load effect

Besides the proportion of correct responses, in order to reveal the underlying processes of the memory load effect in the SDT framework, we also calculated the sensitivity (d') and decision criterion (C). The magnitude of d' reflects how well participants perceptually discriminated between targets and lures. The sign of the C indicates the bias of participants' decisions, with a negative sign corresponding to the response bias toward answering "Yes" (targets) and a positive sign corresponding to the response bias toward answering "No" (lures) (Macmillan & Creelman, 2005).

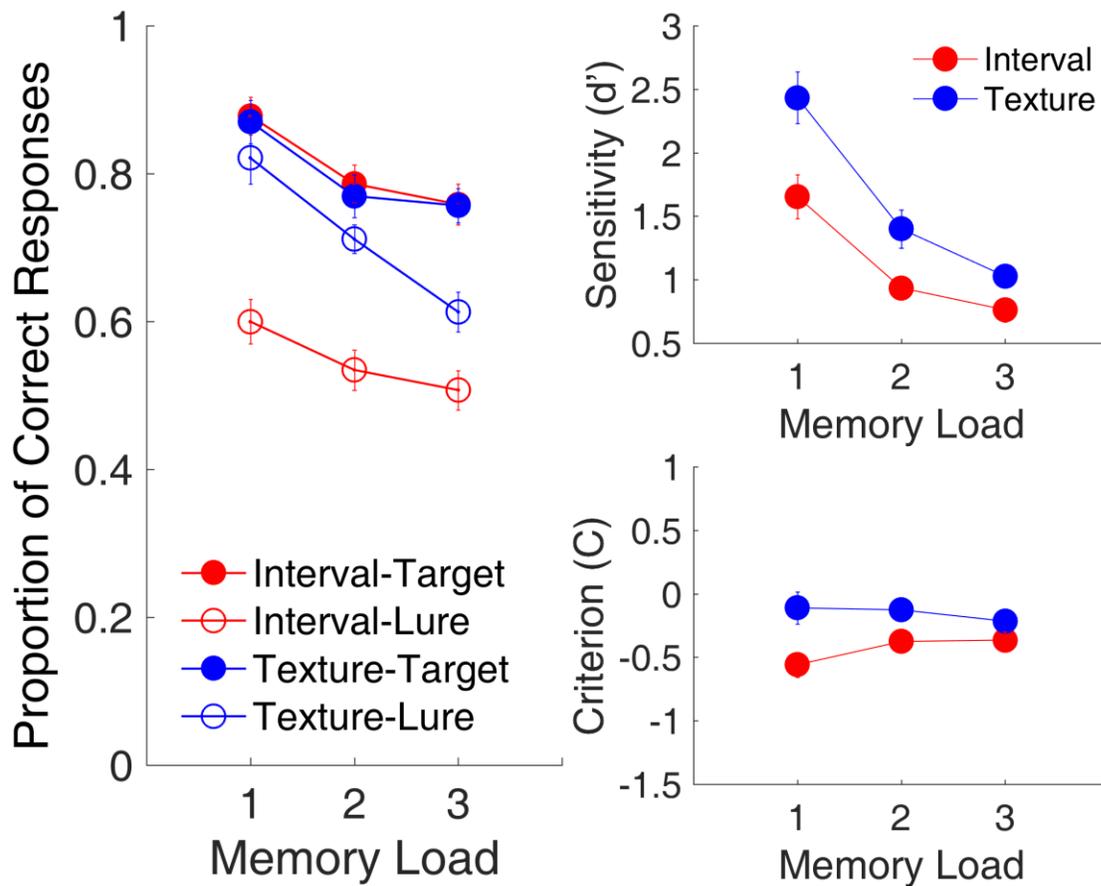


Figure 2.3 Memory load effects in Experiment 1. (A) Proportion of correct responses as a function of memory load. Filled and open circles represent target and lure trials, respectively. (B) d' , as a function of memory load. (C) C , as a function of memory load. Red and blue represent interval and texture trials, respectively. Error bars represent ± 1 standard error of the mean.

Figure 2.3 shows the effects of memory load on the proportion of correct responses, d' , and C plotted across three memory loads and collapsed across serial positions. A repeated-measures ANOVA with three factors (memory load, modality (time intervals or visual textures), and trial type (lure or target)) was conducted on the proportion of correct responses. Results revealed the main effects of memory load ($F(2, 30) = 77.30, p < 0.001, \eta_p^2 = 0.838$, Greenhouse-Geisser [GG]-corrected for non-sphericity), modality ($F(1, 15) = 55.21, p < 0.001, \eta_p^2 = 0.723$, GG-corrected), and trial type ($F(1, 15) = 39.17, p < 0.001, \eta_p^2 = 0.786$, GG-corrected); and a significant interaction

of modality and trial type ($F(1, 15) = 21.72, p < 0.001, \eta_p^2 = 0.592$, GG-corrected). The main effect of memory load indicated that memory performance declined significantly as memory load increased. The interaction of modality and trial type revealed that the differences between interval memory and texture memory differed between the targets and the lures.

Further, a repeated-measures ANOVA with two factors (memory load and modality) on the proportion of correct responses showed only a main effect of memory load for the target trials ($F(2, 30) = 24.68, p < 0.001, \eta_p^2 = 0.622$). For the lure trials, in addition to a main effect of memory load ($F(2, 30) = 17.90, p < 0.001, \eta_p^2 = 0.544$), there was a main effect of modality ($F(1, 15) = 73.55, p < 0.001, \eta_p^2 = 0.831$) and a significant interaction between memory load and modality ($F(2, 30) = 5.10, p = 0.012, \eta_p^2 = 0.254$). The main effect of modality for lure trials, that is, the proportion of correct responses, was significantly lower for lure intervals than for lure textures, suggesting that participants tended to misjudge the lure intervals as targets more often than they tended to misjudge the lure textures as targets.

A repeated-measures ANOVA with two factors (memory load and modality) on d' showed main effects of memory load ($F(2, 30) = 89.72, p < 0.001, \eta_p^2 = 0.857$, GG-corrected) and modality ($F(1, 15) = 25.72, p < 0.001, \eta_p^2 = 0.632$, GG-corrected), and a significant interaction of memory load and modality ($F(2, 30) = 4.61, p = 0.032, \eta_p^2 = 0.235$, GG-corrected). The main effect of modality indicated that, despite our intention to equalize task difficulties between the interval and texture sessions, d' was smaller for time intervals than for visual textures, indicating that the discriminability between targets and lures was better in the texture task than it was in the interval task. The same ANOVA on C showed a main effect of modality ($F(1, 15) = 20.30, p < 0.001, \eta_p^2 = 0.575$, GG-corrected) and a significant interaction of memory load and modality ($F(2, 30) = 7.91, p = 0.005, \eta_p^2 = 0.345$, GG-corrected). The main effect of modality indicated that C was more negative for time intervals, suggesting that participants were more biased when they judged time intervals as targets than when they so judged visual textures. This finding was consistent with the lower proportion of correct responses for the lure trials.

2.2.2.2 *Serial position effect*

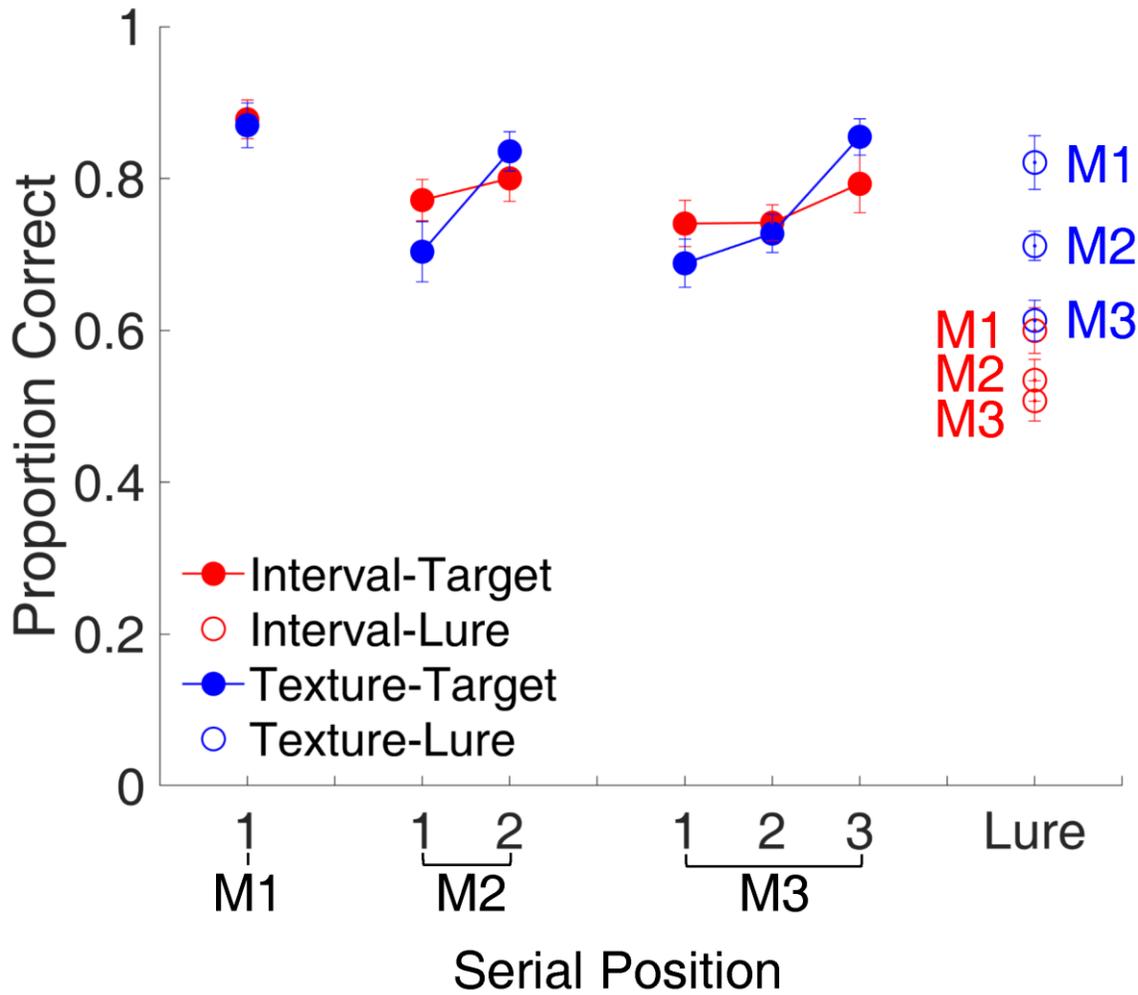


Figure 2.4. Proportion of correct responses as a function of serial position in Experiment 1. M1, M2, and M3 represent the trials within memory load (or the number of study items) of 1, 2, and 3, respectively. Red and blue represent the trials within memory load (or the number of study items) of 1, 2, and 3, respectively. Red and blue represent interval and texture trials, respectively. Filled and open circles represent target and lure trials, respectively. Error bars represent ± 1 standard error of the mean.

Figure 2.4 shows the proportion of correct responses plotted separately for each memory load. The serial position effect was examined only for the target trials. For the lure trials, the proportion of correct responses was plotted by collapsing them across memory loads. The memory load of one study item comprised only one serial

position. No significant difference was observed in the proportion of correct responses between time intervals and visual textures ($t(15) = 0.25, p = 0.808, \text{Cohen's } d = 0.062$). The memory load of two study items comprised Serial Positions 1 and 2. A repeated-measures ANOVA with two factors (serial position and modality) revealed a main effect of serial position ($F(1, 15) = 10.68, p = 0.005, \eta_p^2 = 0.416$). The serial position effect differed between time intervals and visual textures, reflected by a significant interaction between serial position and modality ($F(1, 15) = 8.81, p = 0.010, \eta_p^2 = 0.370$). Further analyses showed a recency effect only for visual textures, that is, the proportion of correct responses at Serial Position 2 was higher than that at Serial Position 1 ($t(15) = 3.88, p = 0.002, \text{Cohen's } d = 0.971$). The memory load of three study items comprised Serial Positions 1, 2, and 3. A repeated-measures ANOVA with two factors (serial position and modality) was conducted, and results indicated a main effect of serial position ($F(2, 30) = 22.20, p < 0.001, \eta_p^2 = 0.597$) and a significant interaction between serial position and modality ($F(2, 30) = 4.69, p = 0.017, \eta_p^2 = 0.238$), again showing that the serial position effect differed between the two modalities. Similarly, there was a recency effect between Serial Positions 2 and 3 for visual textures ($t(15) = 6.73, p < 0.001, \text{Cohen's } d = 1.681, \text{Bonferroni-corrected}$), but the same was not observed for time intervals.

For RTs, only a recency effect was found for visual textures between Serial Positions 2 and 3 in the memory load with three study items ($t(15) = 2.99, p = 0.028, \text{Cohen's } d = 0.747, \text{Bonferroni-corrected}$). In addition, RT was around 0.3 s slower for time intervals than for visual textures ($F(1, 15) = 14.71, p = 0.002, \eta_p^2 = 0.495, \text{GG-corrected}$) across all positions, as revealed by a repeated-measures ANOVA with two factors (serial position and modality).

2.2.2.3 *Similarity effect*

Because similarity increases monotonically with decreasing distance (Kahana & Sekuler, 2002), summed similarity can be determined by the summed probe-study items' distance (PSD), that is, the summation of the 2D Euclidean distances between the probe and all of the study items (Yotsumoro et al., 2008). The summed similarity is inversely proportional to the summed PSD. To evaluate similarity in trials across different memory loads, the summed PSD was divided by the number of study items to derive the mean PSD. Only lure trials were analyzed for the similarity effect, because it is in lure trials that the probe could be similar to the study items, whereas, in target trials, the probe was identical to one of the study items. All lure trials were sorted based on their mean PSDs. Half of

the trials with larger mean PSDs were categorized as small similarity trials, and the other half of the trials with smaller mean PSD were categorized as large similarity trials. It was expected that smaller similarity would lead to more correct rejections of lures as compared with that caused by larger similarity.

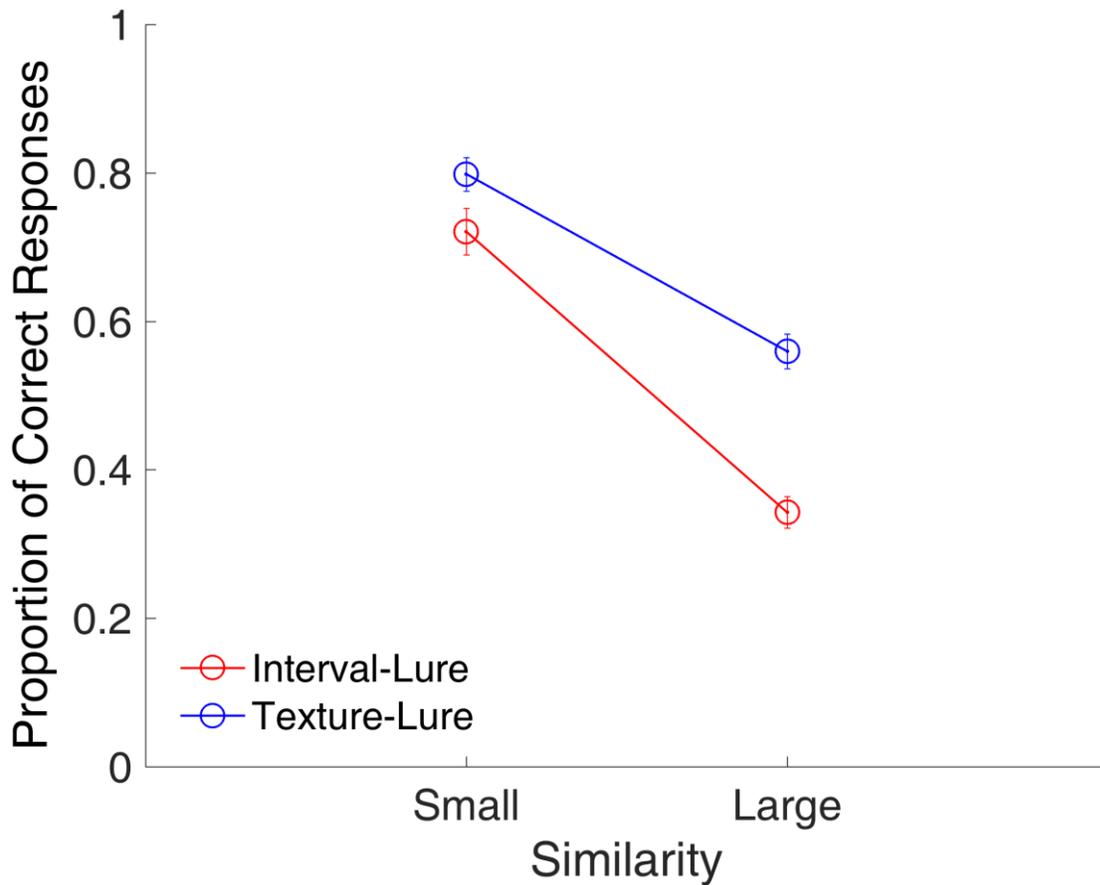


Figure 2.5. Proportion of correct responses as a function of similarity in Experiment 1. Red and blue represent interval and texture trials, respectively. Error bars represent ± 1 standard error of the mean.

Figure 2.5 shows the proportion of correct responses plotted across similarities. A repeated-measures ANOVA with two factors (similarity and modality) indicated main effects of similarity ($F(1, 15) = 252.53, p < 0.001, \eta_p^2 = 0.944$) and modality ($F(1, 15) = 63.25, p < 0.001, \eta_p^2 = 0.808$), and a significant interaction between similarity and modality ($F(1, 15) = 20.80, p < 0.001, \eta_p^2 = 0.581$). The main effect of similarity showed that the

proportion of correct responses was significantly higher for the small similarity condition than for the large similarity condition. The interaction between similarity and modality suggested that the effect of similarity was different between the time intervals and visual textures. Further analysis showed that effect of similarity was larger for time intervals than for visual textures ($t(15) = 4.56, p < 0.001, \text{Cohen's } d = 1.141$).

2.2.2.4 *Binding effect*

In the texture task, the time interval served as the context feature. If there were no binding between the two modalities, the repetition of the interval would not affect the recognition of the visual texture. Similarly, in the interval task, the visual texture served as the context feature. If there were no binding between the two modalities, the repetition of the visual texture would not affect the recognition of the time interval either. We conducted a post-hoc analysis to evaluate this binding effect.

Figure 2.6 shows the proportion of correct responses plotted across the repetition and switch of the context features. We conducted a repeated-measures ANOVA with three factors (context, modality, and trial type). Results indicated main effects of modality ($F(1, 15) = 48.23, p < 0.001, \eta_p^2 = 0.763$, GG-corrected) and trial type ($F(1, 15) = 38.32, p < 0.001, \eta_p^2 = 0.719$, GG-corrected), a significant interaction of modality and trial type ($F(1, 15) = 17.01, p < 0.001, \eta_p^2 = 0.531$, GG-corrected), and a marginally significant three-way interaction ($F(1, 15) = 4.26, p = 0.057, \eta_p^2 = 0.221$, GG-corrected). This finding of no main effect of context but a marginal three-way interaction of context, modality, and trial type suggested that the effects of context might differ among the combinations of modality and trial type. Further analysis with a t-test only revealed a marginal effect of the repetition of the temporal features on lure textures ($t(15) = -1.95, p = 0.070, \text{Cohen's } d = 0.489$). This implies that the repetition or switching of context features might affect the rejection of the lures in different ways for time intervals and visual textures. Further, the same analysis on RTs showed no effect of context (all $ps > 0.147$) for either targets or lures in the two modalities.

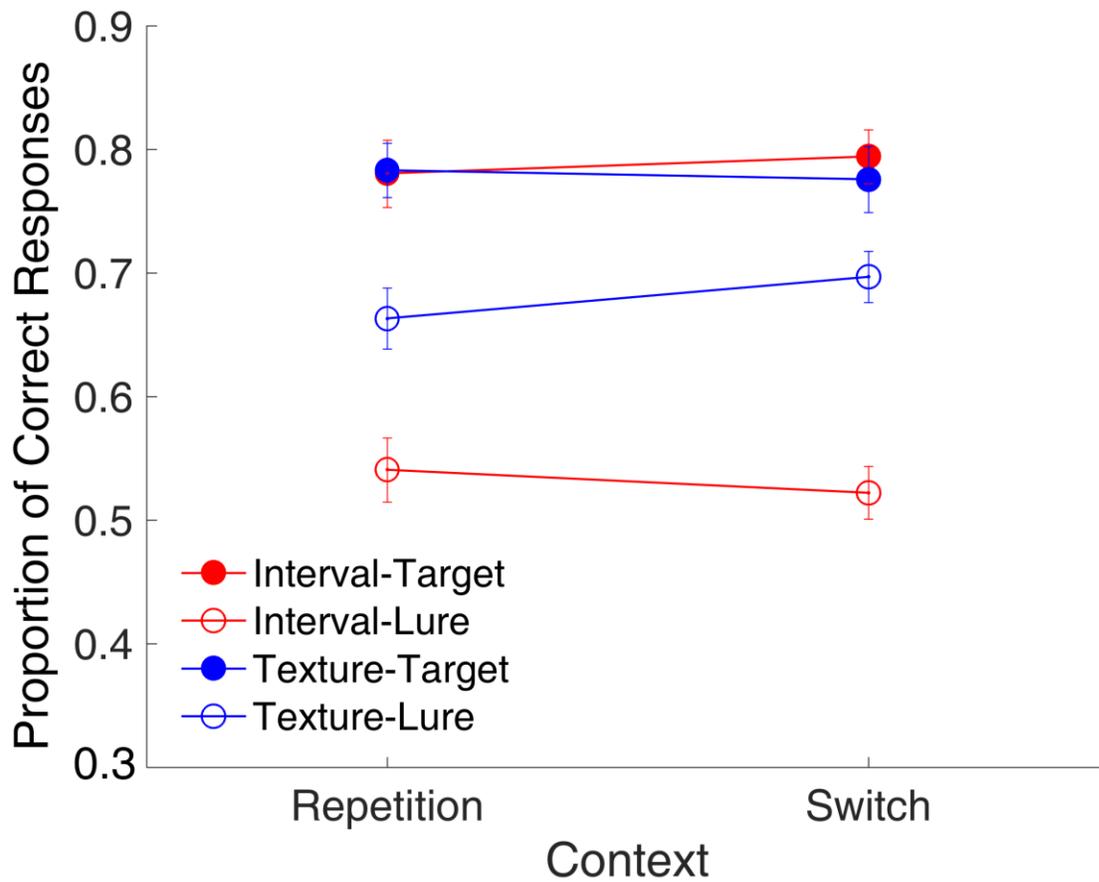


Figure 2.6. Effects of the context on the proportion of correct responses in Experiment 1. Red and blue represent interval and texture trials, respectively. Filled and open circles represent target and lure trials, respectively. Error bars represent ± 1 standard error of the mean.

2.2.3 Discussion

Using Sternberg’s item recognition task, we revealed the similar working memory characteristics for time intervals and for visual textures, specifically the effects of memory load and similarity, supporting the hypothesis that time intervals can be represented in working memory as discrete items similarly to visual objects. However, the serial position effects differed between interval and texture memory. The serial position effect was only found for visual textures but not for time intervals. In addition, we did not find strong effects of binding.

The memory load effect for time intervals showed a very similar pattern to that for visual textures. This result indicated that working memory for time intervals of visual events is also affected by memory load, as it is for those of auditory events (Manohar & Husain, 2016; Teki & Griffiths, 2014). This is consistent with the memory load effect for time intervals in other paradigms like absolute identification (Lacouture, Grondin, & Mori, 2001).

The similarity effect was observed for both time intervals and visual textures. By simply dividing the lure trials into the two categories of large similarity (small mean PSD) and small similarity (large mean PSD) trials, we found a similar pattern of similarity effect for both time intervals and visual textures. In Experiment 2, we further investigated the similarity effect by systematically manipulating similarities.

However, the serial position effect, specifically the recency effect, was found only for visual textures. The lack of the primacy effect could be attributed to the short ISIs between the presentations of items (Glenberg et al., 1980). No recency effect was found for time intervals. From the pattern of the data, however, a weak serial position effect for time intervals was observed. Because Experiment 1 contained three memory loads, the number of trials in each memory load condition was constrained, which may not have been adequate for revealing the possible weak recency effect for time intervals. In the next experiments, we investigated the serial position effect in detail.

In addition, we did not find strong evidence for the binding of temporal features and visual features. No effect of the repetition of the visual textures was found on interval memory. On the other hand, there seemed to be a weak effect of the repetition of the temporal features on the rejections of lure textures. Because the studies of feature binding between temporal features and non-temporal features are limited (Bogon et al., 2017), the underlying mechanism remains unclear. One explanation for the weak or absent binding effect may be that task-irrelevant non-spatial features (as temporal and visual features in this study) are weakly bound or not bound to the same events (Chen, 2009; Hommel, 1998). Another explanation might be that feature binding is affected by learning, through which the over-learned feature combinations in long-term memory may facilitate the binding of the same combinations of features (Colzato, Raffone, & Hommel, 2006; Hommel & Colzato, 2009). In Experiment 2, we simplified the visual textures and further examined the effects of binding.

We also found that participants were more biased to judge lure intervals as targets than to so judge lure textures, which might lead to the lower proportion of correct responses for lure intervals. In addition, the average RT for time intervals was about 0.3 s larger than that for visual textures. This may be because time is always estimated

after visual events (Bogon et al., 2017). In this context, evidence suggests that time intervals may be a high-level stimulus feature encoded at a later stage in the visual processing hierarchy (Maarseveen, Hogendoorn, Verstraten, & Paffen, 2017).

2.3 Study 1: Experiment 2

In Experiment 2, we further investigated the effects of serial position, as well as similarity and binding, on working memory for time intervals and visual textures, using only the three-item memory load condition. The SDT was incorporated into the recognition task to further investigate the characteristics of the interval memory.

2.3.1 Materials and Methods

2.3.1.1 *Participants*

Twelve naïve volunteers from the University of Tokyo (six females, mean age: 25 years, range: 19–32 years) participated in Experiment 2, after excluding three participants because the average accuracy of their recognition of either visual textures or time intervals was at around the chance level, and for one participant, the mean RT for visual textures was 2 standard deviations away from the mean RT across all participants.

2.3.1.2 *Stimuli and Apparatus*

Similar to Experiment 1, 2D compound sinusoidal gratings were used as visual stimuli. Only nine gratings were chosen from the 25 gratings that were used in Experiment 1, with relatively large Euclidian distances from each other, such as 1.51, 1.51; 1.51, 1.75; 2, 1.51; 1.75, 2; 2, 2; 2.25, 2; 2, 2.49; 2.49, 2.25; and 2.49, 2.49 CPD, combined with 9 time intervals derived from a logarithmically spaced range of 0.2512–0.9772 s. The logarithmically spaced setting for time intervals was used to increase the absolute distance between long time intervals so that the distinguishability between long and short time intervals were balanced.

In Experiment 2, three gratings were randomly assigned to three out of nine different spatial frequencies, and three out of nine different time intervals, as study items. The items were labeled as units 1 to 9 according to their Euclidean distances from the origin of coordinates in the spatial frequency ((0, 0) CPD) or in the time interval (0 s). One unit PSD meant the distance between two neighbor units, and 2 units PSD meant the distance between a unit

and the unit next to its neighbor, and so on. A probe grating had the following constraints: the PSD between the probe and the target study item was zero, and the PSD between the probe and the non-target study items was at least two units, for either spatial frequencies or time intervals. In addition, the maximum PSD was set to five units, to avoid the location of time intervals at two ends (that is, the shortest or longest ones). The probe was also set such that it was not equally closest to more than one study item. Finally, we constrained summed PSD to six, seven, or eight units for target trials, and 9, 10, or 11 units for lure trials. The same sets of stimuli were used in the interval and texture sessions. The apparatus used in Experiment 2 was the same as that used in Experiment 1.

2.3.1.3 Procedure

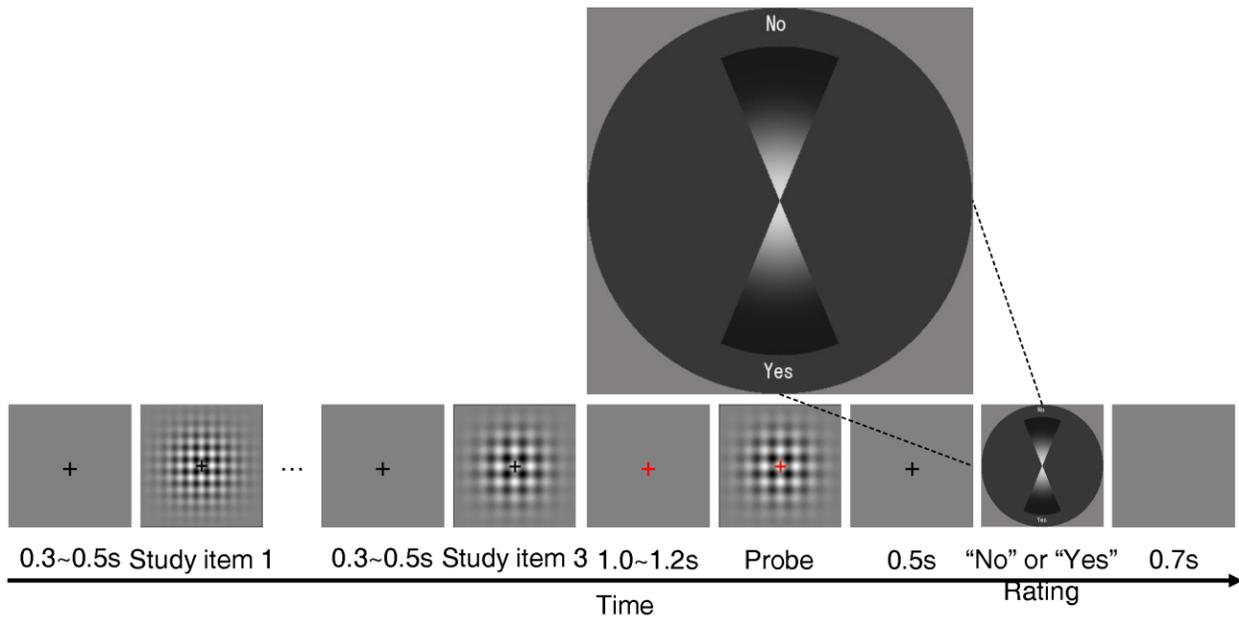


Figure 2.7. Time course of the recognition/working memory task in Experiment 2. The enlarged disk with two vertical sectors represents the panel for “Yes” and “No” judgment and confidence rating by mouse clicking. Participants responded “Yes” by clicking on the lower sector, and responded “No” by clicking on the upper sector. The location of the mouse click corresponded to their confidence: participants were instructed to click further away from the center if they were more confident but to click closer to the center if they were not very confident.

The procedure for this experiment was the same as for Experiment 1, except that there were always three study items, and the responses were registered using a rating panel (Figure 2.7). A rating panel with a dark gray panel and two vertical sectors representing “Yes” and “No” judgments were presented 0.5 s after the probe disappeared. The participants judged whether the probe was the same as one of the study items and rated their confidence in the judgment on the rating panel. Half of the participants responded “Yes” by clicking on the lower sector, and responded “No” by clicking on the upper sector. The location of the mouse click corresponded to their confidence: the participants were instructed to click further away from the center if they were more confident but to click closer to the center if they were not very confident. The position of the “Yes” and “No” panels were reversed for the other half of the participants. Participants were asked to respond as quickly as possible, without sacrificing accuracy.

Each session consisted of 18 practice trials and 324 test trials composed of a combination of factors such as serial position (1, 2, or 3), similarity (large, medium, or small), context feature (repetition or switch), and trial type (target trials or lure trials).

2.3.2 Results

The average accuracy across all available participants for the time intervals was $57.47\% \pm 3.96\%$ and for the visual textures was $65.26\% \pm 4.76\%$; the average RT across all available participants for the time intervals was 1.89 ± 0.51 s and for the visual textures was 1.68 ± 0.28 s.

2.3.2.1 Receiver operating characteristic curve

Figure 2.8 shows the average ROC curves for the time intervals and visual textures. Despite our intention to equalize task difficulties between the interval session and the texture session, the ROC curves indicated that sensitivity (d') was smaller for time intervals than for visual textures, indicating that the discriminability between targets and lures was better in the texture task than in the interval task. The curvatures were similar between time intervals and visual textures, which suggested that the shapes of the distributions of the decision criterion (C) for the targets and lures were similar for the two modalities.

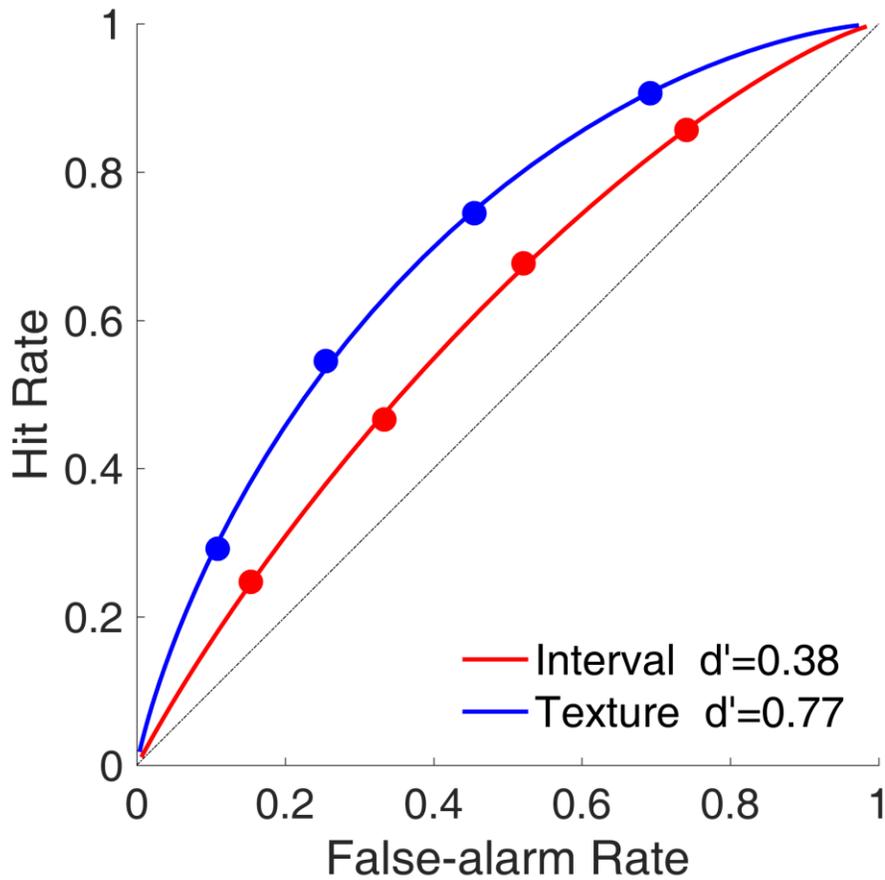


Figure 2.8. Receiver operating characteristic curves (hit rate as a function of false-alarm rate) in Experiment 2. Red and blue represent interval and texture trials, respectively.

2.3.2.2 *Serial position effect*

Figure 2.9 shows the proportion of correct responses plotted across the three serial positions. A repeated-measures ANOVA with two factors (serial position and modality) indicated a main effect of serial position ($F(2, 22) = 14.53, p < 0.001, \eta_p^2 = 0.569$). The serial position effect differed between the time intervals and visual textures, with a marginal significance ($F(2, 22) = 3.38, p = 0.053, \eta_p^2 = 0.235$). Further analysis indicated the presence of a last-item recency effect for visual textures as observed in Experiment 1, that is, there was a significant difference between Positions 2 and 3 ($t(11) = -3.06, p = 0.032, \text{Cohen's } d = 0.884, \text{Bonferroni-corrected}$). For time intervals, no significant difference was found between Positions 2 and 3, but there was a significant difference between Positions

1 and 2 ($t(11) = -2.85, p = 0.047$, Cohen's $d = 0.823$, Bonferroni-corrected). Further, we observed a significant difference between time intervals and visual textures for the lure trials ($t(11) = -5.05, p < 0.001$, Cohen's $d = 1.458$), suggesting that participants were more biased toward judging lure intervals as targets. The ANOVA on RT revealed no significant effects on RT (all $ps > 0.1$).

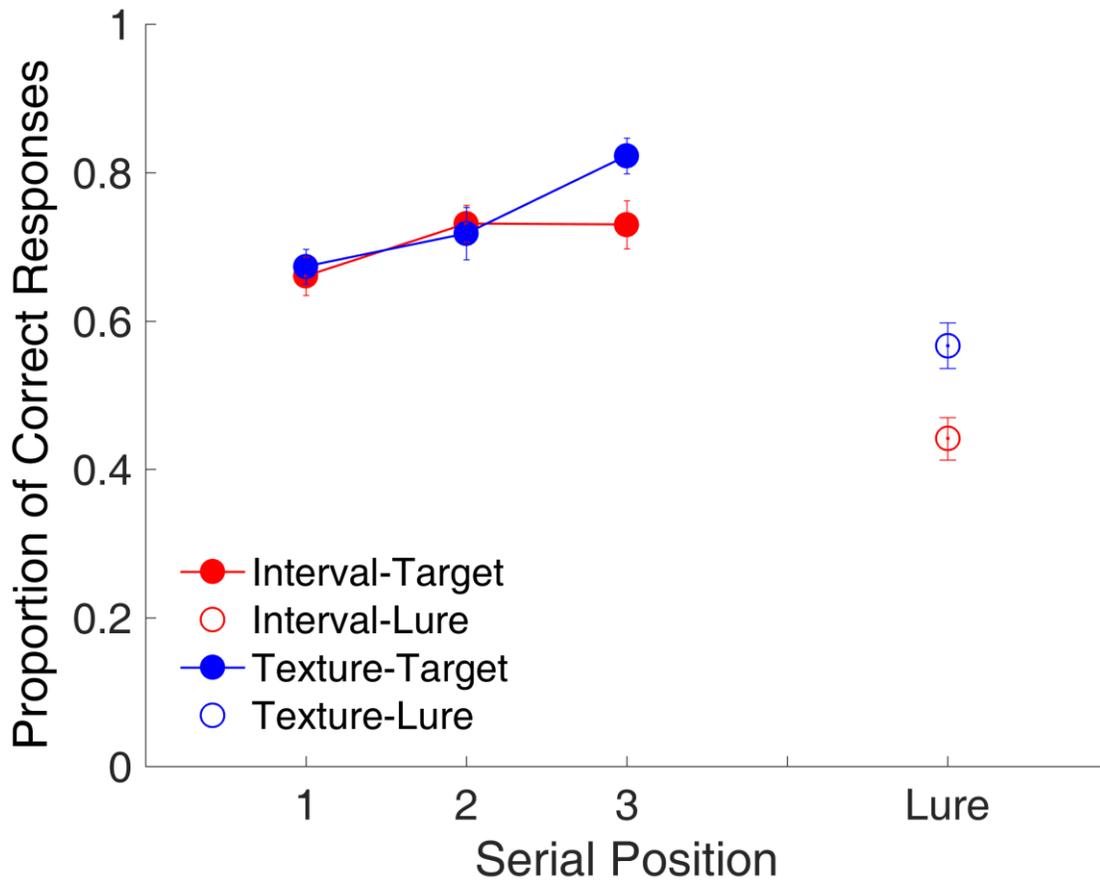


Figure 2.9. Proportion of correct responses as a function of serial position in Experiment 2. Red and blue represent interval and texture trials, respectively. Filled and open circles represent target and lure trials, respectively. Error bars represent ± 1 standard error of the mean.

2.3.2.3 *Similarity effect*

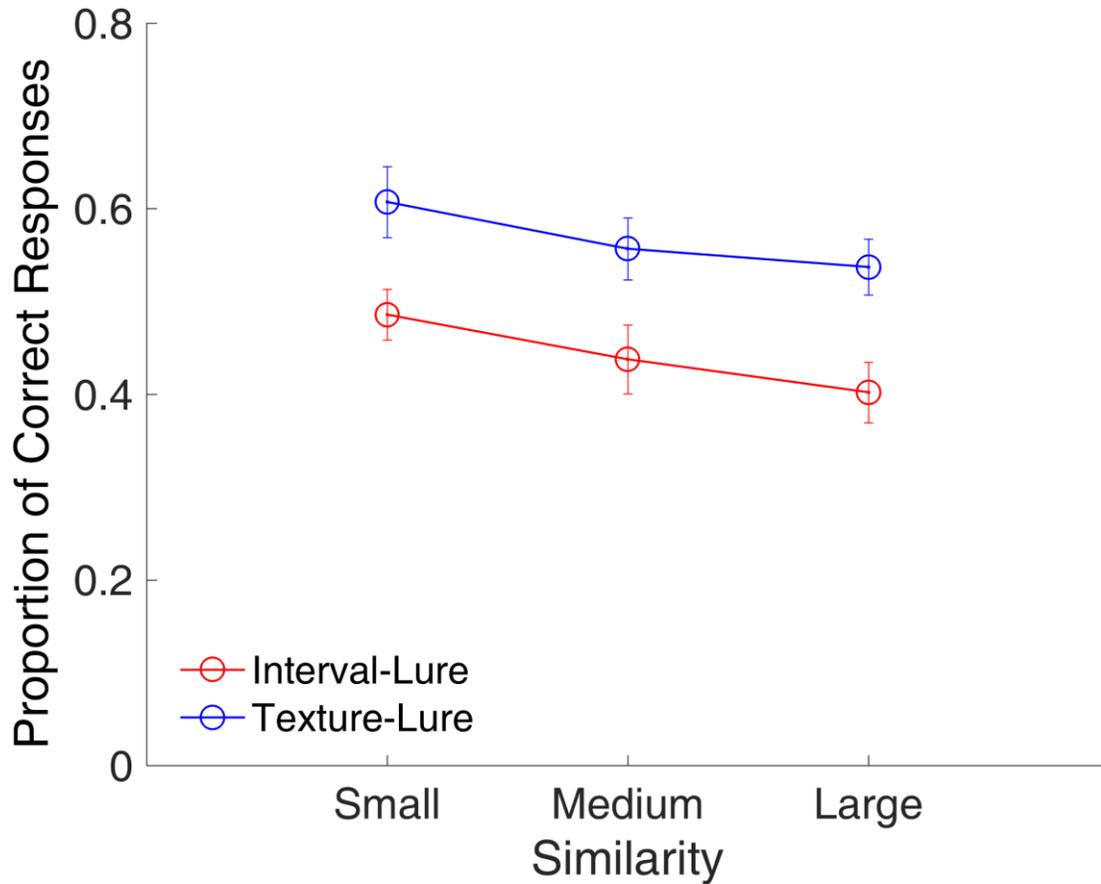


Figure 2.10. Proportion of correct responses as a function of similarity in Experiment 2. Red and blue represent interval and texture trials, respectively. Error bars represent ± 1 standard error of the mean.

Figure 2.10 shows the proportion of correct responses plotted across the small, medium, and large similarities. A repeated-measures ANOVA with two factors (similarity and modality) indicated a main effect of similarity ($F(2, 22) = 13.31, p < 0.001, \eta_p^2 = 0.548$) and modality ($F(1, 11) = 25.32, p < 0.001, \eta_p^2 = 0.697$), indicating that the participants were more likely to consider a lure interval as a target as compared to a lure texture.

2.3.2.4 Binding effect

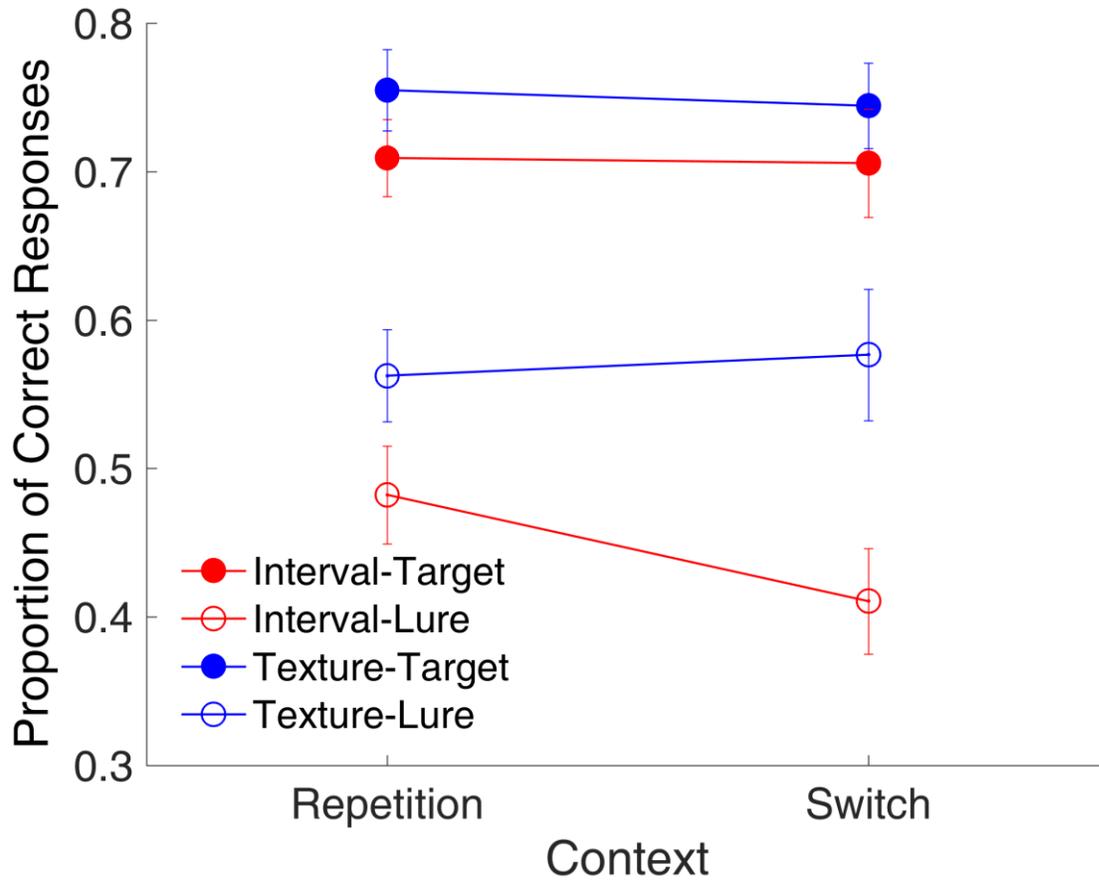


Figure 2.11. Proportion of correct responses as a function of the context in Experiment 2. Red and blue represent interval and texture trials, respectively. Filled and open circles represent target and lure trials, respectively. Error bars represent ± 1 standard error of the mean.

Figure 2.11 shows the proportion of correct responses plotted across the repetition and switch of the context features. Again, we conducted a post-hoc analysis of the binding effect. Unlike Experiment 1, in Experiment 2, it was possible that, when both the temporal feature and visual feature of the probe matched the features of the study items, the two features did not belong to the same study item. Therefore, we selected trials in which the two features belonged to the same study item. We conducted a repeated-measures ANOVA with three factors (context, modality, and trial type). Results indicated main effects of modality ($F(1, 11) = 15.64, p = 0.002, \eta_p^2 = 0.587$) and trial type

($F(1, 11) = 29.37, p < 0.001, \eta_p^2 = 0.728$). Although there were no main effects of context and no significant interactions between context and other factors, further analysis with a t-test indicated a significant difference between the repetition and switching of context textures for the lure intervals (Figure 2.11, $t(11) = 2.24, p = 0.0467$, Cohen's $d = 0.647$). The same analysis on RTs showed no effects of context (all $ps > 0.592$).

2.3.3 Discussion

By combining the recognition task with a signal detection task, we provided further evidence to support the idea that time intervals can be stored in and retrieved from working memory as discrete items. The ROC curves showed similar curvatures between time intervals and visual textures, suggesting similar distributions of C for judging temporal and visual items. The serial position effect was found for time intervals, a robust similarity effect was observed again, but a strong binding effect was still not found. We also confirmed that lures in the interval memory were more likely to be misperceived as targets.

Although we did not observe a clear last-item recency effect for time intervals, we still found the serial position effect for the interval memory. According to a two-component account (Allen, Baddeley, & Hitch, 2014; Phillips & Christie, 1977), the absence of a last-item recency effect may be attributed to the interference of the last item by subsequent events. In the present experiment, the rating panel may have disrupted the recency effect for time intervals because participants had to make a relatively complicated decision regarding rating their confidence during the judgment. The absence of a recency effect was also seen in some conditions in a study by Manohar and Husain (2016). They claimed that recency effects are susceptible to attentional disruption by the presence of irrelevant information.

After carefully manipulating the similarity between probe and study items, we observed a robust similarity effect, as also observed in Experiment 1, for both time intervals and visual textures. These results confirm that time intervals can be stored as discrete but noisy items, similarly to visual working memory, and that working memory for temporal items is also influenced by temporal similarities between the items.

2.4 Study 1: Experiment 3

In Experiment 3, we further examined the memory characteristics of the longer duration of the supra-second range. We used the same recognition task as that used in Experiment 1, with only three study items, to examine the characteristics of working memory only for time intervals in the sub- with the supra-second range. We mixed and randomized the trials in the two ranges in each session. In this experiment, we only measured interval memory.

2.4.1 Materials and Methods

2.4.1.1 *Participants*

Ten naïve volunteers from The University of Tokyo (three males, mean age: 20 years, range: 19–21 years) participated in Experiment 3, after excluding two participants because their overall accuracy for both gratings and time intervals was around the chance level.

2.4.1.2 *Stimuli and apparatus*

2D gratings were presented with different time intervals in either a sub- or supra-second range. The spatial frequencies of the 2D gratings were at the same Euclidian distance from the origin as (2, 2) CPD, but with five different radians from a linear spaced range of $\pi/2 \times 5/16$ to $\pi/2 \times 11/16$. Subsequently, the gratings were of 2D spatial frequencies such as 1.33, 2.49; 1.68, 2.27; 2, 2; 2.27, 1.68; and 2.49, 1.33 CPD. In this way, only the phases (the apparent texture) but not the distances were different, reducing the variability between visual textures.

For the time intervals, we first randomly selected 15 intervals from the sub-second range of 0.5–0.7 s and another 15 from the supra-second range of 2–2.8 s. We then derived a set of five intervals by adding $0.2 \times (-2, -1, 0, 1, 2)$ s to each of the selected intervals. The target study item was the middle interval after adding the 0. The other two non-target study items were selected from the remaining four intervals. In the target trials, the probe matched the target study item, and in the lure trials, the probe was selected from the two remaining unselected intervals.

The apparatus used in Experiment 3 was the same as that used in Experiments 1 and 2.

2.4.1.3 *Procedure*

The procedure of Experiment 3 was the same as that of Experiment 1, except for the use of the memory load of three study items only. Each participant completed one session with mixed trials in either the sub- or supra-second range. The time intervals were in the same range within each trial. There were 18 practice trials and 360 test trials composed of a combination of factors such as time range (sub- or supra-second), serial position (1, 2, or 3), context (repetition or switch), and trial type (target trials or lure trials).

2.4.2 **Results**

The average accuracy across all available participants for the sub-second intervals was 61.68%±4.45% and for the supra-second intervals was 63.54%±4.36%; the average RT across all available participants for the sub-second intervals was 0.79±0.32 s and for the supra-second intervals was 0.86±0.36 s.

2.4.2.1 *Serial position effect*

Figure 2.12 shows the proportion of correct responses plotted across serial positions. A repeated-measures ANOVA with two factors (serial position and modality) indicated no main effect of serial position ($F(2, 18) = 1.22, p = 0.319, \eta_p^2 = 0.119$), but a main effect of time range ($F(1, 9) = 12.63, p = 0.006, \eta_p^2 = 0.584$), indicating that the proportion of correct responses was significantly higher for the sub- than for the supra-second interval. The ANOVA on RTs indicated that the RT was significantly lower ($F(1, 9) = 9.59, p = 0.013, \eta_p^2 = 0.516$) for the sub- than for the supra-second interval.

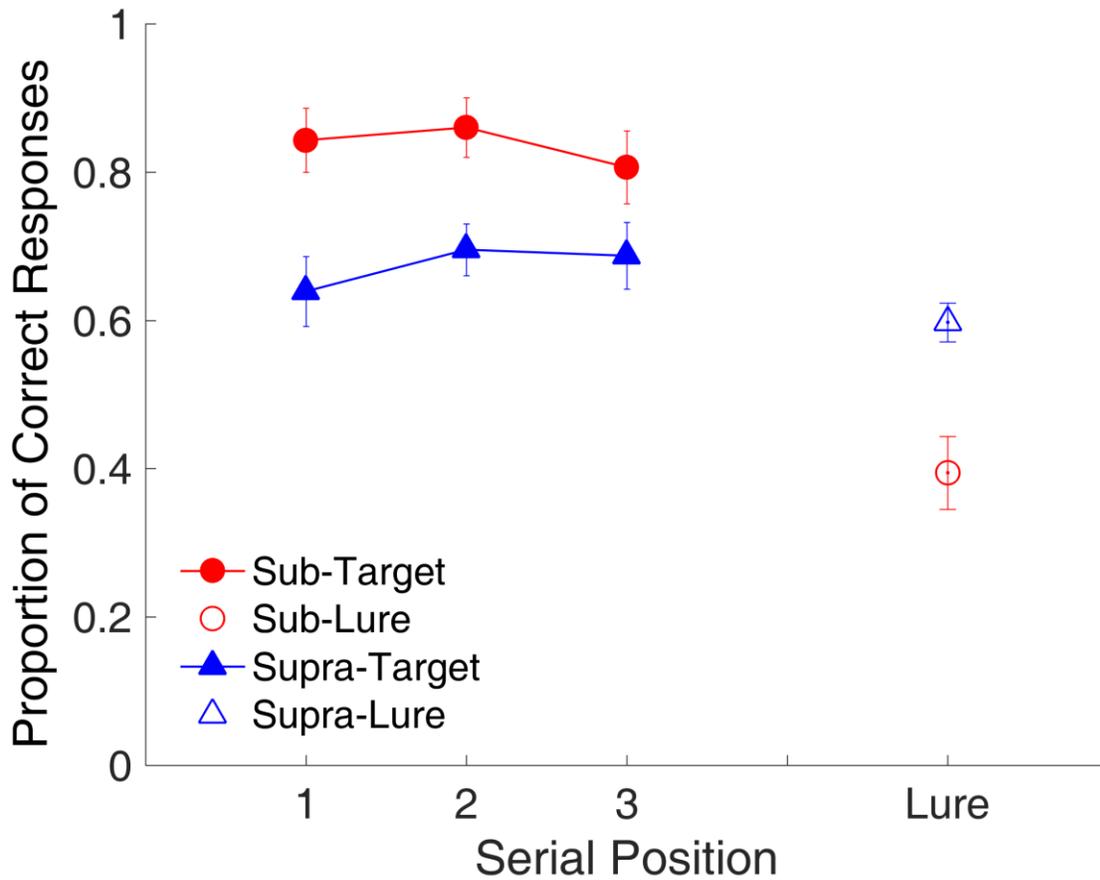


Figure 2.12. Proportion of correct responses as a function of serial position in Experiment 3. Red and blue represent sub-second and supra-second trials, respectively. Filled circles and triangles represent target trials, and open circles and triangles represent lure trials. Error bars represent ± 1 standard error of the mean.

2.4.2.2 Similarity effect

To compare the stages of processing of between the sub- and supra-second intervals, we examined how similarity affected the d' and C pertaining to the “Yes/No” judgment.

Figure 2.13 shows the effects of similarity on the proportion of correct responses, d' , and C . We conducted a repeated-measures ANOVA with two factors (similarity and modality) for each performance measure. The results indicated robust similarity effects on the proportion of correct responses ($F(2, 18) = 24.79, p < 0.001, \eta_p^2 = 0.734$), d' ($F(2, 18) = 16.07, p < 0.001, \eta_p^2 = 0.641$, GG-corrected), and C ($F(2, 18) = 10.88, p < 0.001, \eta_p^2 = 0.547$). Further,

we observed significant effects of time range on the proportion of correct responses ($F(1, 9) = 20.21, p = 0.001, \eta_p^2 = 0.692$) and C ($F(1, 9) = 19.68, p = 0.002, \eta_p^2 = 0.686$), indicating that participants were more likely to judge lures in the sub-second intervals as targets than to judge lures in the supra-second intervals as targets. The difference in decision criterion (C) instead of in the sensitivity (d') between the two time ranges suggests that the difference between the processing of the sub- and of supra-second intervals might occur in the decision-making stage.

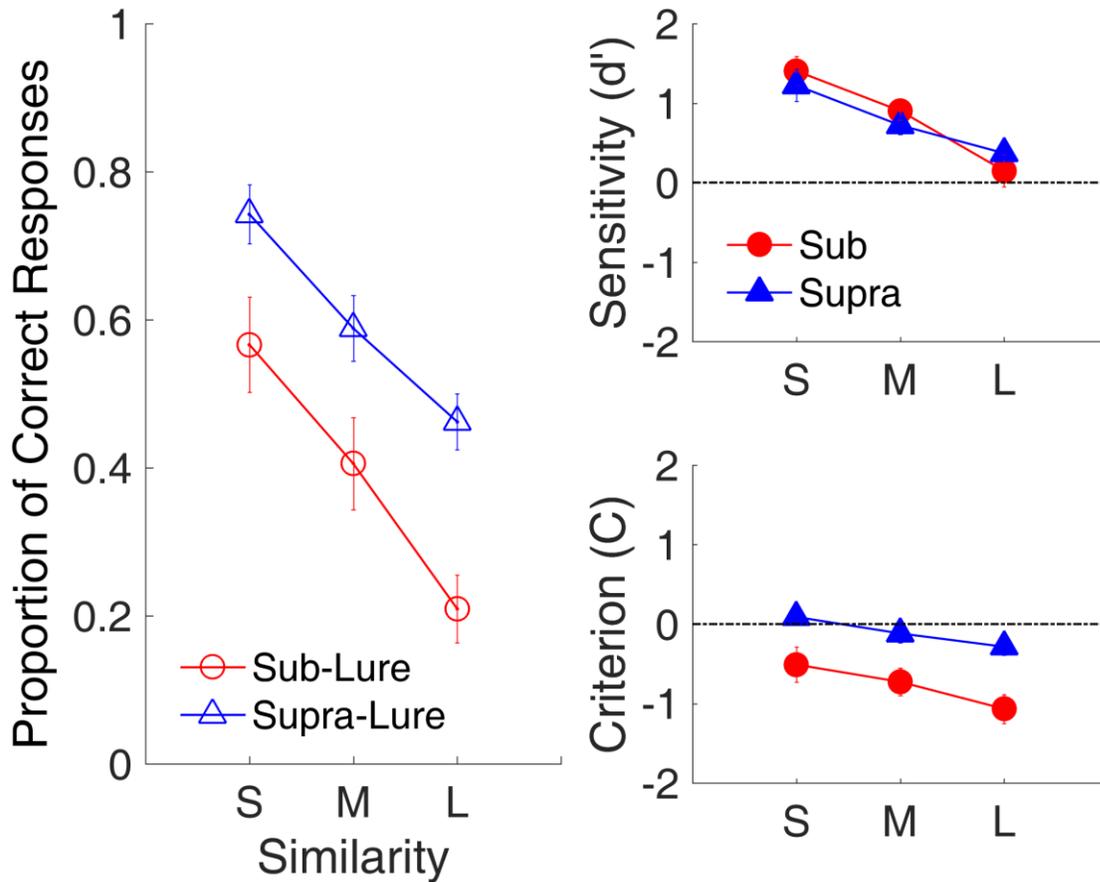


Figure 2.13. Similarity effect in Experiment 3. (A) Proportion of correct responses as a function of similarity. Open circles and triangles show lure trials. (B) d' as a function of similarity. (C) C as a function of similarity. Red and blue represent sub-second and supra-second trials, respectively. Error bars represent ± 1 standard error of the mean.

2.4.2.3 *Binding effect*

Figure 2.14 shows the effect of context on the proportion of correct responses. A repeated-measures ANOVA with three factors (context, time range, and trial type) was conducted. Results indicated a main effect of trial type ($F(1, 9) = 17.89, p = 0.002, \eta_p^2 = 0.665$, GG-corrected); two-way significant interactions of context and time range ($F(1, 9) = 6.80, p = 0.028, \eta_p^2 = 0.431$, GG-corrected), and of time range and trial type ($F(1, 9) = 20.95, p = 0.001, \eta_p^2 = 0.700$, GG-corrected); and a significant three-way interaction of context, time range, and trial type ($F(1, 9) = 5.57, p = 0.043, \eta_p^2 = 0.382$, GG-corrected).

Similar to the binding effect observed in Experiment 1, the significant three-way interaction suggested that the effects of context differed among the combinations of modality and trial type. Further analysis with a t-test showed a marginal difference between the repetition and switch of the context ($t(9) = 2.22, p = 0.054$, Cohen's $d = 0.703$) for the lure sub-second intervals, consistent with the findings of Experiment 2. This result suggested that the repetition of visual textures facilitated the correct rejection of the lure intervals; however, for supra-second intervals, the repetition increased the false-alarm rate ($t(9) = -3.30, p = 0.009$, Cohen's $d = 0.845$).

The same analysis on RT showed the main effect of trial type ($F(1, 9) = 22.44, p = 0.001, \eta_p^2 = 0.714$, GG-corrected); two-way significant interactions of context and time range ($F(1, 9) = 13.80, p = 0.005, \eta_p^2 = 0.605$, GG-corrected), of context and trial type ($F(1, 9) = 9.41, p = 0.013, \eta_p^2 = 0.511$, GG-corrected), and of time range and trial type ($F(1, 9) = 6.76, p = 0.029, \eta_p^2 = 0.429$, GG-corrected). The repetition of the visual textures facilitated recognition of the target sub-second intervals by reducing the RT ($t(9) = -2.74, p = 0.023$, Cohen's $d = 0.866$); for supra-second intervals, however, the repetition increased the RT ($t(9) = 9.03, p < 0.001$, Cohen's $d = 0.720$).

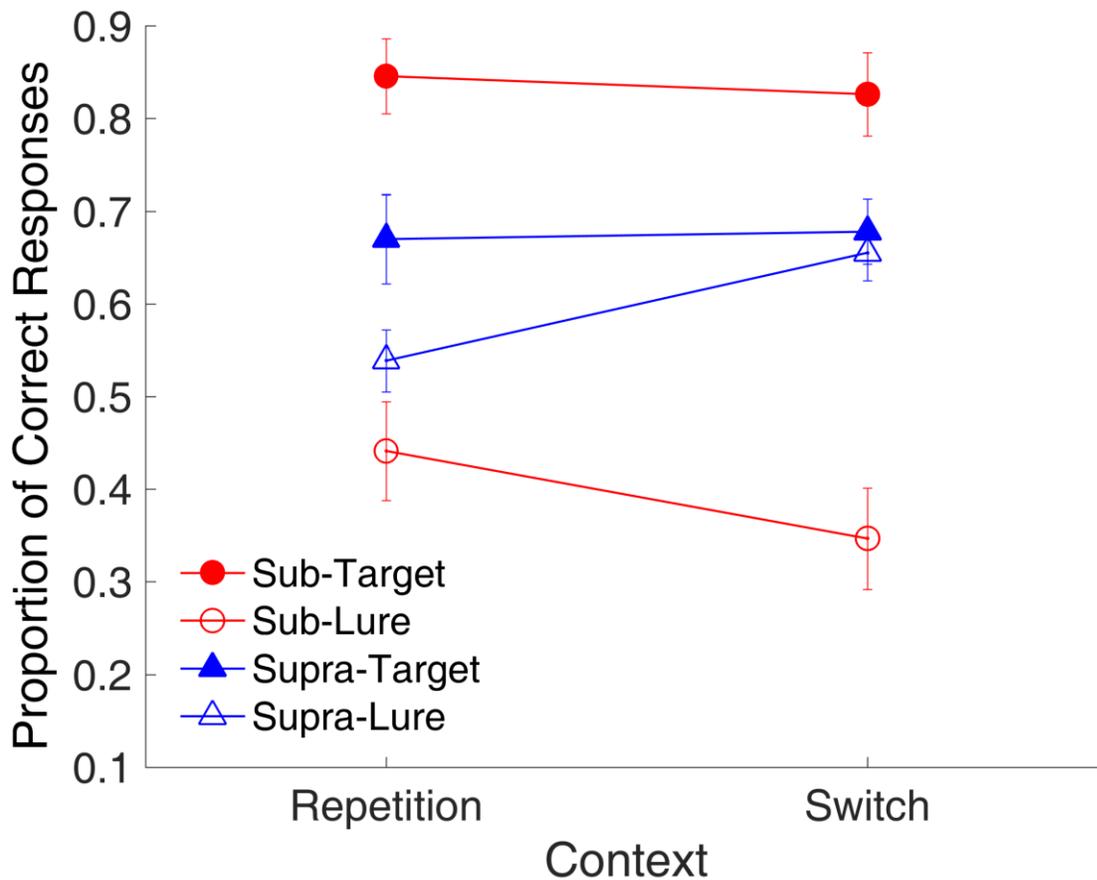


Figure 2.14. Effects of context on the proportion of correct responses in Experiment 3. Red and blue represent sub-second and supra-second trials, respectively. Filled circles and triangles show target trials, and open circles and triangles show lure trials. Error bars represent ± 1 standard error of the mean.

2.4.3 Discussion

Using the temporal recognition/working memory task, we compared the effects of serial position, similarity, and binding for time intervals in the sub-second range with those in the supra-second range. The effect of similarity was observed for both ranges. Specifically, the repetition of context features had opposite effects between the two ranges, but no serial position effect was found for either range of time intervals.

The absence of the serial position effect for supra-second intervals was consistent with the findings of Teki and Griffiths (2014). On the other hand, for sub-second intervals, the present finding was not consistent with the

results of Experiment 2, where the serial position effect was observed. One reason might be the mixing of the supra- and sub-second intervals in a single session. Specifically, the brain network recruited for the processing of supra-second intervals in previous trials may affect the processing of subsequent peri-second intervals (Murai & Yotsumoto, 2016a). Therefore, the mixing might have led to an interference with the processing of peri-/sub-second intervals such that it disrupted the serial position effect for sub-second intervals. It is difficult to draw strong conclusions from the present evidence, and the absence of recency effects must be interpreted with caution. Further research is necessary to clarify this issue.

The difference of processing of between sub- and supra-second intervals may occur in the later stage of processing. Previous studies on the similarity effect showed that participants are less likely to judge a probe as one of the study items when the similarity among the study items are high (Kahana & Sekuler, 2002; Nosofsky & Kantner, 2006; Viswanathan, Perl, Visscher, Kahana, & Sekuler, 2010). In our experiment, the supra-second intervals may be less distinctive, as they are perceived less precisely (Murai & Yotsumoto, 2016b; but see Lewis & Miall, 2009) as compared to sub-second intervals. This predicts that sub-second intervals are more likely to be judged as the targets, as was the case in the present study. The difference in this bias might occur in the decision-making stage instead of during sensory processing, as reflected by the difference in the decision criterion (C) instead of the sensitivity (d') between the sub- and supra-second intervals.

Further, the difference in the proportion of correct responses between the repetition and switching of the context features was found only for the lure intervals in both ranges, but in the opposite way. Specifically, the repetition of visual textures led to a higher correct rejection of lure sub-second intervals but a lower correct rejection for lure supra-second intervals, as compared with that caused by the switching of visual textures. If there is indeed a binding and the retrieval of previous bindings occurs, according to Bogon et al. (2017), the partial repetition should lead to worse recognition, and it was so for the lure trials in the supra-second intervals. However, the same was not observed for sub-second intervals. No context effect was found for target intervals, that is, the repetition/switch of visual textures did not seem to contribute to the recognition of target intervals.

2.5 Study 1: General discussion

Using Sternberg's item recognition task and framework of SDT, we examined working memory for multiple time intervals and compared its characteristics, including the effects of memory load, serial position, and similarity, with that for visual textures. In Experiment 1, we observed the same patterns of memory load effect and similarity effect between time intervals (in sub-second range) and visual textures, supporting the hypothesis that similarly to visual objects, time intervals are represented in working memory as discrete items. However, the serial position effect was absent for the interval memory. In Experiment 2, we further investigated serial position effect using a rating scale. We found a serial position effect for the time intervals, although it was weak and slightly different from that for visual textures, further supporting our hypothesis. In Experiment 3, we examined interval memory of the supra-second range, and found the supra-second intervals might also be represented as discrete items, although there were differences between working memory for the sub- and that for the supra- second intervals.

2.5.1 Item representations of time intervals in working memory

Working memory for multiple time intervals can be compared to working memory for multiple visual objects. Visual working memory has limited storage, and memory performance is influenced by the number of study items, which yields a memory load effect (Bays & Husain, 2008; Luck & Vogel, 1997). The memory load effect for time intervals showed a similar pattern as that for visual textures, especially for targets that had already been stored in working memory. This finding suggests that similar to visual textures, time intervals can be represented as discrete items in working memory, and the countable number of temporal items influences the so-called temporal working memory. The memory load effect for time intervals has also been observed in previous studies with auditory stimuli (Lacouture et al., 2001; Manohar & Husain, 2016; Teki & Griffiths, 2014). However, these studies did not directly compare temporal working memory with auditory working memory. Therefore, future studies should make direct comparisons and confirm whether the working memory representation of time intervals can be similar to that of auditory or other non-temporal items, as to that of visual items.

According to the noisy exemplar model (Kahana & Sekuler, 2002), visual stimuli are represented as separate but noisy items in working memory, and the recognition of visual items is influenced by the similarity between the probe and study items; time intervals may be represented in the same way. Our results showed the same

patterns of the similarity effect for lure intervals as for lure textures, and thus confirmed that the item representations of time intervals were similar to those of visual textures. The similarity effect for visual items has been well investigated in previous studies (for example, Kahana & Sekuler, 2002; Yotsumoto et al., 2008; Zhou, Kahana, & Sekuler, 2004), but that for temporal items was examined for the first time in the present study.

The noisy-item representations of time intervals are further supported by the findings pertaining to the ROC curve. A regular and symmetric ROC curve implied the probability distributions of the decision criterion (C) for targets and lures. The similar curvature of the ROC curves found in our study indicates that the C distributions for time intervals share common properties with those for visual textures. Although the C distributions for target intervals and lure intervals may overlap, target intervals can be distinguished from lure intervals according to a specific C. The noisy exemplar model proposed that the recognition memory of items is also affected by the inter-study item similarity (Kahana & Sekuler, 2002). Therefore, future studies need to examine this proposal with reference to temporal items.

The item representations of time intervals are of some difference from those of visual objects. Although we found in Experiment 2 that the later presented intervals were recognized better than were those presented earlier, consistent with the definition of the recency effect, this recency effect was different from the last-item recency effect observed for visual textures. Further, we observed the absence of the serial position effect for sub-second intervals in Experiment 1 and 3, and for supra-second intervals in Experiment 3. Traditionally, there are two possible accounts for the recency effects for visual items. The first is the two-component (Allen et al., 2014; Phillips & Christie, 1977) based dual-store model, which proposes that the later presented items are stored in an active but limited-capacity buffer. In this account, attention is allocated to the current presented item, and it is switched from the earlier item to the later item. Therefore, the later presented items are retrieved better than the earlier ones are, as they are no longer in the buffer but are instead likely to be in long-term storage. This account cannot explain the weak or even absent recency effect for temporal items. The other account, the distinctiveness-based single-store model (Neath, 1993), stresses the role of temporal distinctiveness between the study items. According to this model, later presented items are relatively more distinctive at retrieval and are therefore retrieved more easily. This account cannot explain the absence of the last-item recency effect for temporal items either.

A possible explanation may be that the item representations of time intervals are weaker than those for visual objects are. In this case, recognition relies more on perceptual averaging across sequential items, which would ignore the serial positions of items, than on the identities of individual items (Ariely, 2001; B. Zhou, Zhang, & Mao, 2015). Another possible explanation may be related to the dual role of time intervals. Time intervals of events are not only the contents per se, but they are also the temporal framework for the contents of events (for example, visual textures in the study) stored in and retrieved from recognition memory (Bogon et al., 2017; Fournier & Gallimore, 2013; Thavabalasingam, O'Neil, Zeng, & Lee, 2016).

Another difference is that lure items in the sub-second range are more likely to be misjudged as target items as compared with lure items in the visual modality. At the same time, the judgment for target items in both modalities yielded very similar performances. It is not clear why the hit rates were similar while the false-alarm rates were discrepant between time intervals and visual textures. In the SDT framework, both the difference of d' and C would lead to these results. Future studies need to control d' between the interval and texture tasks to examine whether this tendency occurred due to C , and thus in the stage of decision-making, or whether this tendency is related to the domain-general working memory processes, regardless of the modality.

2.5.2 Mechanisms underlying the item representations of time intervals

The mechanisms underlying the item representations for the temporal and visual modalities may share common properties. As for visual objects, the neural object-file theory proposes that objects are first individualized to discrete items, named as “object files” (Kahneman, Treisman, & Gibbs, 1992). However, at the first stage of processing, item representations are not sufficient to be recognized, and it is at the second stage of processing that the binding of features into the object files make the objects distinctive, such that they can be recognized as targets or lures (Xu & Chun, 2009). There may be a similar mechanism for the individualization and recognition of time intervals as discrete items in working memory.

Temporal and visual features belonging to the same events might interact with each other because of the binding. Our results suggested the presence of possible binding effects between visual memory and interval memory. In our study, the effects were only seen on lures but not on targets. It may be because the task-irrelevant feature is more weakly bound to the event as compared with the task-relevant feature. Additionally, the repetition of

the temporal feature led to worse recognition performance for the visual features, consistent with that for the auditory feature (Bogon et al., 2017). The mechanism underlying these findings is not yet clear, and therefore, more studies need to examine the binding of temporal and non-temporal features.

2.5.3 Difference between working memory for sub- and supra- second intervals

Time intervals in sub- and supra-second ranges might be retrieved differently. In our study, the difference between the recognition of sub- and supra-second intervals was characterized by the difference in the value of C . Sensitivity (d') was similar for both ranges, meaning that the early processing of the two ranges of time intervals were similar. The difference in C between the sub- and supra-second intervals suggests that the difference in working memory for the two ranges of time intervals may occur in the later stage of processing, like in the decision-making stage. Further, sub-second intervals tended to be judged as the targets more often as compared to supra-second intervals. Previous studies suggested that hierarchical timing networks are recruited for processing time intervals in two ranges, with lower-level processes affecting both ranges but having a stronger effect on sub-second intervals. Further, higher-level processes have been found to govern both ranges but have a stronger effect on supra-second intervals (Murai & Yotsumoto, 2016b; Petter, Lusk, Hesslow, & Meck, 2016; Wiener, Turkeltaub, & Coslett, 2010).

As for the neural mechanisms underlying interval perceptions of sub- and supra- seconds, a number of studies have reported distinct but partially overlapped mechanisms for those two ranges (Murai & Yotsumoto, 2016b; Petter et al., 2016; Wiener et al., 2010). The cerebellum has been reported to be involved in both sub- and supra- second timing (for humans, Petter et al., 2016; for monkey, Ohmae, Kunimatsu, & Tanaka, 2017). Supra-second intervals (for example, >2000ms) have been reported as being further processed by the cortico-thalamic-striatal circuit (Petter et al., 2016). The shared neural mechanisms (for example, cerebellum) could be a part of the shared earlier stage of processing, and the neural networks specific to the supra-second intervals (for example, the cortico-thalamic-striatal circuit) could be the level that differentiates the decision stage (C).

Chapter 3 to Chapter 7 are excluded because these parts are scheduled to be published in the near future.